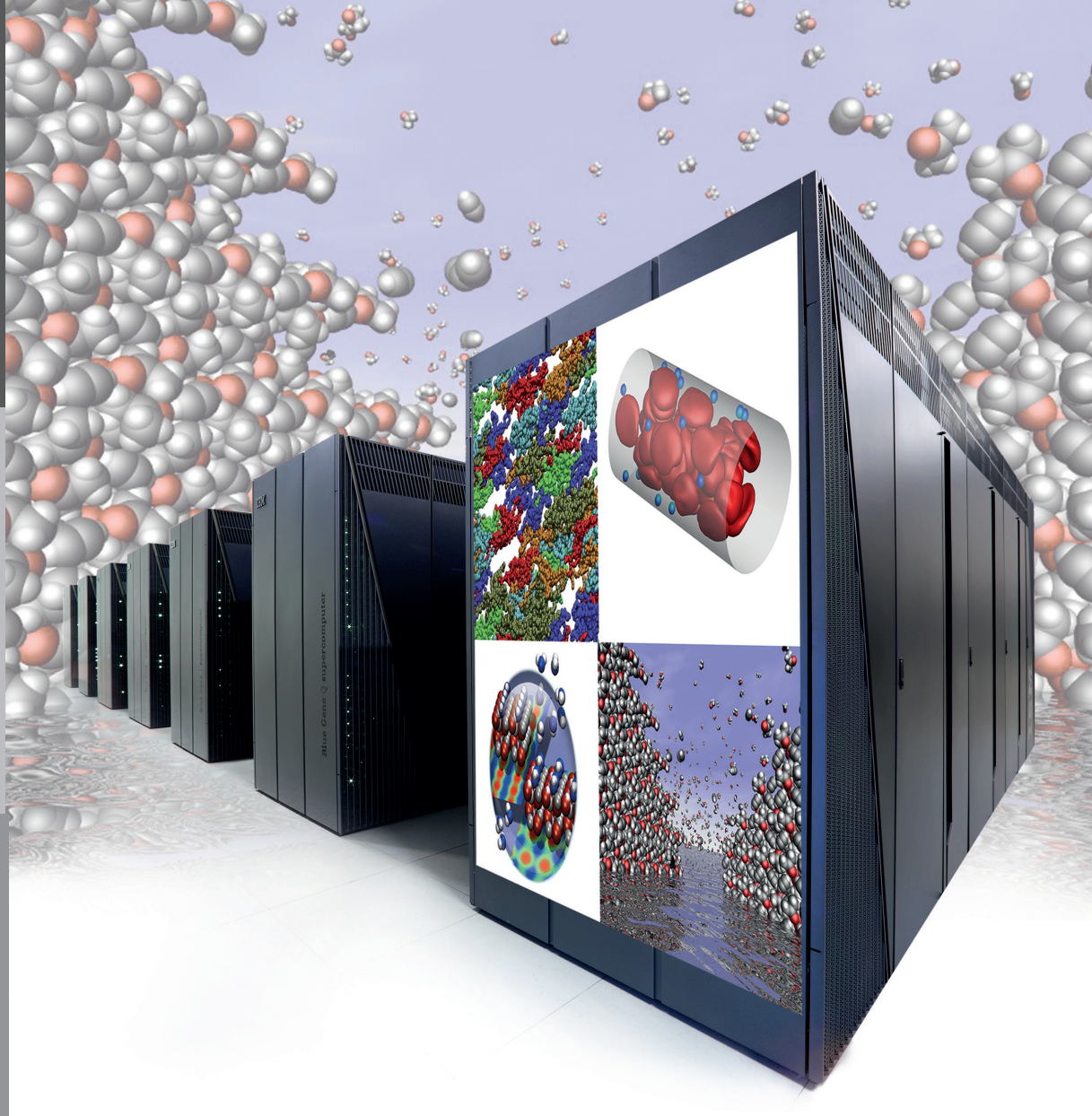


Computational Trends in Solvation and Transport in Liquids
Lecture Notes

edited by Godehard Sutmann, Johannes Grotendorst, Gerhard Gompper, Dominik Marx

cecam
Centre Européen de Calcul Atomique et Moléculaire

JÜLICH
FORSCHUNGSZENTRUM

Member of the Helmholtz Association

Schriften des Forschungszentrums Jülich
IAS Series Volume 28

Forschungszentrum Jülich GmbH
Institute for Advanced Simulation (IAS)
Jülich Supercomputing Centre (JSC)

Computational Trends in Solvation and Transport in Liquids

edited by Godehard Sutmann, Johannes Grotendorst,
Gerhard Gompper, Dominik Marx

Jülich CECAM School, 23 – 27 March 2015
Forschungszentrum Jülich GmbH
Lecture Notes

Schriften des Forschungszentrums Jülich
IAS Series Volume 28

ISSN 1868-8489 ISBN 978-3-95806-030-2

Bibliographic information published by the Deutsche Nationalbibliothek.
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Publisher and Distributor:	Forschungszentrum Jülich GmbH Zentralbibliothek 52425 Jülich Phone +49 (0) 24 61 61-53 68 · Fax +49 (0) 24 61 61-61 03 e-mail: zb-publikation@fz-juelich.de Internet: http://www.fz-juelich.de/zb
Cover Design:	Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH
Printer:	Grafische Medien, Forschungszentrum Jülich GmbH
Copyright:	Forschungszentrum Jülich 2015

Schriften des Forschungszentrums Jülich
IAS Series Volume 28

ISSN 1868-8489
ISBN 978-3-95806-030-2

The complete volume is freely available on the Internet on the Jülicher Open Access Server (JuSER) at www.fz-juelich.de/zb/openaccess

Persistent Identifier: urn:nbn:de:0001-2015020300

Preface

The present Jülich CECAM School on Computational Science, which is co-supported by the Cluster of Excellence Ruhr Explores Solvation RESOLV, has a main focus on "Solvation Science", which is increasingly recognized as an interdisciplinary field akin to "Materials Science" or "Neuroscience". The School addresses a variety of different computational and simulation methods, appropriate for those hierarchies of time- and length-scales which are a challenge for modeling and simulation in Solvation Science. Therefore, the School is intended as a continuation of former Schools, organized in Jülich in 2009 and 2012, focusing on *Multiscale Simulation Methods in Molecular Sciences* and *Hierarchical Methods for Dynamics in Complex Molecular Systems*. The key of the present School is to introduce to a broad scientific audience modern computational methods which originate from a wide background but which can be applied to a variety of different scales, both in time and length, to describe solvation and transport phenomena on different levels of approximation. Problems associated with solvation and transport emerge in various distinct fields, ranging from fundamental questions in wet chemistry or soft matter physics to industrial applications in chemical engineering, which naturally induces a highly interdisciplinary character. Therefore, one aim of this School is to bring people together from diverse disciplines and various background in addition to triggering exchange of ideas between established experts and students - being the next generation of researchers.

More specifically, the Jülich CECAM School 2015 focuses on computational trends, multi-method approaches and modeling in the field of solvation and transport in liquids. It covers the field from large-scale coarse-grain modeling down to fully quantum-mechanical simulations of liquids at the level of electrons and nuclei. Not only bulk liquids and homogeneous solutions will be discussed, but also heterogeneous systems such as liquid/solid interfaces as well as solvated (bio)molecules. In particular, recent advances in adaptive resolution methods both in the realm of finite element modeling and of interfacing atomistic and coarse-grain descriptions of liquids are presented. Moreover, a variety of hybrid methods, such as QM/MM approaches for solvated biomolecules, continuum solvation and Lattice-Boltzmann techniques, will be part of the program. Coarse-graining in many distinct flavors is certainly an important approach to describe transport in liquids and solvation of increasingly complex systems. This includes methods such as Brownian dynamics for biomolecular recognition, mesoscale particle methods for hydrodynamics to model particle-fluid interactions or neural network potentials to describe water not just as an inert but as a potentially reactive solvent.

The recent revival of liquid-state integral equation methods in combination with molecular dynamics and electronic structure theory is acknowledged in the program as well as new developments in molecular density functional theory of aqueous solutions. Well-established techniques such as force field molecular dynamics for large-scale simulations and ab initio molecular dynamics for wet chemical reactions will be addressed as well in order to introduce the audience to these fields. Within molecular dynamics special atten-

tion will be given to methods that are tailored to unravel solvation effects. Presentation of most recent developments in path integral simulation techniques will complement the lectures to cover the smallest scales down to the nuclear level, which are relevant in the atomistic description of solvation.

The solution of large-scale complex problems needs a direct link to high performance computing, which today includes the use of GPUs in addition to massively parallel CPU based systems. Recent developments and trends will be addressed in the School not only by providing lectures but also by including a hands-on practical tutorial on elementary GPU programming.

The target group of this Jülich CECAM School on Computational Science are young scientists, especially PhD students and early postdocs.

Many individuals have significantly contributed to the success of the School and we are particularly grateful to all the lecturers who have prepared and contributed the lecture notes, which form the basis of the present book that will be made available at the School. Without their effort, compiling such an excellent reference book on *Computational Trends in Solvation and Transport in Liquids* would not have been possible.

We are greatly indebted to the School's secretaries Helga Offergeld and Britta Hoßfeld, who were indispensable for this School by taking care of logistics, transport, registration and catering. Very special thanks go to Monika Marx and Ilona Lütje, who realized technically the present Lecture Notes as well as the poster booklet by not only collecting but also carefully editing the written contributions. Last but not least, technical support during the lectures and hands-on sessions by Oliver Bücker and René Halver is greatfully acknowledged.

Jülich and Bochum
March 2015

Godehard Sutmann
Johannes Grotendorst
Gerhard Gompper
Dominik Marx

Contents

Atomistic Methods

Coarse Grain and Continuum Methods

Hybrid Methods

Mesoscale Fluid Methods

Numerical Methods and Hardware

Modeling the Quantum Nature of Atomic Nuclei by Imaginary Time Path Integrals and Colored Noise

Michele Ceriotti

Laboratory of Computational Science and Modeling, IMX,
École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

E-mail: michele.ceriotti@epfl.ch

The vast majority of atomistic simulations of matter treat the nuclei as classical point particles, evolving in time according to Hamiltonian dynamics, and following classical Boltzmann statistics. As a matter of fact, this is quite a harsh approximation when dealing with hydrogen-containing compounds. Hydrogen nuclei – as well as other light elements – exhibit significant deviations from classical behavior up to and above room temperature. Here I will present an overview of path integral methods, that can be used to include nuclear quantum effects in atomic-scale models. I will cover the basic theory, discuss some of the practicalities in the implementation and use, and finally introduce the most recent advances towards making these techniques less computationally demanding by using colored (correlated) stochastic dynamics.

Simulations that describe matter at the level of individual atoms offer a very high level of accuracy, transferability and predictive power. In order to make them practically feasible, a number of approximations are often introduced that trade off the accuracy in describing some physical effects in exchange for a reduced complexity and computational cost. Perhaps the most widely adopted approximation is the decoupling of the electronic structure problem from that of the statistical and dynamical behavior of the atomic nuclei. This can take the form of Born-Oppenheimer approximation[1] – where the ground-state electronic structure problem is solved for a given configuration of the nuclei – or can be realized by modeling the interaction between the atoms using an empirical force field that represent effectively (and inexpensively) the potential energy surface for the atoms treated as point particles.

The Born-Oppenheimer approximation is generally very satisfactory, except when the system evolves in an electronic excited state or for a few ultra-fast chemical reactions. However, it only consists in a factorization of the combined electronic-nuclear wavefunction, and in principle the nuclei should be treated as quantum particles. The vast majority of atomistic simulations are performed with an additional approximation, that is to treat the nuclei as classical particles that evolve in time following Hamilton's equations and that are subject to Boltzmann, classical statistics. These are certainly reasonable approximations at high temperature, and when dealing with heavy nuclei. If however one compares the thermal energy $k_B T$ and the quantum of harmonic energy $\hbar\omega$ for a molecular vibration of frequency ω at temperature T, it will become clear that for many compounds $\hbar\omega/k_B T \gg 1$ even well above room temperature, which casts some shadows on the consequences of neglecting the quantum nature of the nuclear degrees of freedom in simulations.

There are several examples of the impact of the quantum mechanical behavior of nuclei on experimental observables. The heat capacity of substances deviates from the Dulong-Petit prediction of $3k_B T$ per atom (that corresponds to classical statistics for a harmonic crystal), in particular for stiff bonds (as in diamond) or for hydrogen-containing com-

pounds. The kinetic energy distribution of atomic nuclei, as measured by neutron Compton scattering, differs dramatically from the Maxwell-Boltzmann distribution[2]. Reaction rates at low temperature do not follow an Arrhenius behavior. The stability of different compounds or phases varies with isotope composition, and one can for instance estimate (extrapolating the values measured for ^1H_2O, ^2H_2O, ^3H_2O) that the pH of water would be around 8.5 if nuclei behaved classically. Some of these phenomena – isotope substitution effects in particular – simply cannot be observed in the absence of a quantum mechanical treatment of the nuclear degrees of freedom, while others entail a deviation of computed properties from their experimental counterparts. Neglecting nuclear quantum effects (NQEs) is particularly detrimental in the case of *ab initio* molecular dynamics, where the ground-state electronic structure problem is solved on the fly, and the nuclei evolve on the bare Born-Oppenhemier potential energy surface. Simulations employing empirical force fields can include NQEs indirectly, by fitting the parameters of the inter-atomic potential to experimental observables, or more directly using approximate techniques such as Feynman-Hibbs effective potentials[3].

Solving the Schrödinger equation for the nuclei is impractical except for very simple systems. Here we will discuss how the imaginary-time path integral formalism[4–7] can be used to evaluate accurately NQEs in complex condensed-phase applications[8–10]. We will focus on static, equilibrium properties, but will briefly mention extensions to the path integral formalism that can be used to treat approximately quantum dynamics[11,12]. We will focus on the case in which different nuclei can be treated as distinguishable particles, which is often true except for cases at cryogenic temperatures. Particle exchange statistics can be included within a path integral formalism, but at the cost of considerable complication and an increase of the computational cost[5,13].

1 Imaginary-Time Path Integrals

The path integral formulation of quantum mechanics makes it possible to express all the quantities that describe a physical system in terms of exponential averages of an appropriate action integral over the possible paths joining two points in phase space – much like the minimum action principle makes it possible to formulate classical mechanics as the minimization of the action over a tentative path[4]. Furthermore, it makes it possible to express the quantum mechanical partition function at inverse temperature $\beta = 1/k_B T$ [a]

$$Z = \mathrm{Tr}\, e^{-\beta \hat{H}}$$

as the *path integral*

$$Z = \oint \mathcal{D}\left[q\left(\tau\right)\right] e^{-\frac{1}{\hbar}\int_0^{\beta\hbar}\left[\frac{1}{2}m\dot{q}(\tau)^2 + V(q(\tau))\right]\mathrm{d}\tau}. \tag{1}$$

The symbol $\oint \mathcal{D}\left[q\left(\tau\right)\right]\cdot$ is a functional integral over all the possible *closed* paths in configuration space, weighed with the exponential of an action-like integral over the path. In the following we will discuss how to give a practical definition of Eq. (1), and how to use this formalism to compute experimental observables including NQEs.

[a] We consider for simplicity the case of a single particle with position q and mass m in an external potential V

1.1 Trotter Factorization

Start by writing the partition function in the position representation

$$Z = \int \mathrm{d}q_1 \left\langle q_1 \middle| e^{-\beta \hat{H}} \middle| q_1 \right\rangle .$$

The Hamiltonian can be written as the sum of a potential and kinetic energy terms, $\hat{H} = \hat{V} + \hat{T}$, and the position ket is an eigenstate of the potential energy, so that $e^{-\beta \hat{V}} |q_1\rangle = e^{-\beta V(q_1)} |q_1\rangle$. Unfortunately, one cannot factor $e^{-\beta \hat{H}}$ into the product $e^{-\beta \hat{V}} e^{-\beta \hat{T}}$, because potential and kinetic energy are not commuting operators. However, the error in doing such a factorization decreases when $\beta \hat{H}$ becomes small. So, one could write[b]

$$e^{-\beta \hat{H}} = \left(e^{-\beta \hat{H}/P} \right)^P \approx \left(e^{-\beta_P \hat{V}/2} e^{-\beta_P \hat{T}} e^{-\beta_P \hat{V}/2} \right)^P + \mathcal{O}\left(\beta_P^2\right),$$

which becomes exact in the $P \to \infty$ limit. Note that we have also introduced the shorthand $\beta_P = \beta/P$. One can show that the partition function converges to the exact quantum mechanical result with a leading error of $\mathcal{O}\left(\beta^2/P^2\right)$, and in practice for a system with a maximum frequency $\omega_{\max}$ one needs a number of imaginary time slices that is at least a small multiple of $\beta\hbar\omega_{\max}$.

One can then introduce $P-1$ closure relations $\int \mathrm{d}q_j \, |q_j\rangle \langle q_j|$, obtaining

$$\begin{aligned} Z \approx Z_P = \int \mathrm{d}q_1 \ldots \mathrm{d}q_P \Big[& \left\langle q_1 \middle| e^{-\beta_P V(q_1)/2} e^{-\beta_P \hat{T}} e^{-\beta_P V(q_2)/2} \middle| q_2 \right\rangle \ldots \\ & \ldots \left\langle q_P \middle| e^{-\beta_P V(q_P)/2} e^{-\beta_P \hat{T}} e^{-\beta_P V(q_1)/2} \middle| q_1 \right\rangle \Big] . \end{aligned} \tag{2}$$

The terms with the potential energy are just scalar values, that can be brought outside the quantum mechanical brackets. One is then left with a series of terms corresponding to the off-diagonal elements of the kinetic energy operator, that are readily evaluated by transforming in the momentum representation:

$$\begin{aligned} \left\langle q_i \middle| e^{-\beta_P \hat{T}} \middle| q_j \right\rangle &= \int \mathrm{d}p \left\langle q_i \middle| e^{-\beta_P \hat{T}} \middle| p \right\rangle \langle p | q_j \rangle = \\ &= \frac{1}{2\pi\hbar} \int \mathrm{d}p e^{-\beta_P p^2/2m} e^{\mathrm{i}p(q_i - q_j)/\hbar} = \\ &= \frac{1}{2\pi\hbar} \sqrt{\frac{2\pi m}{\beta_P}} e^{-\frac{1}{2}\beta_P m \omega_P^2 (q_i - q_j)^2} \end{aligned} \tag{3}$$

where we have used $\langle p|q\rangle = e^{-\mathrm{i}pq/\hbar}/\sqrt{2\pi\hbar}$, performed the integral over the momentum and introduced the spring constant $\omega_P = 1/\beta_P\hbar$. Plugging Eq. (3) into Eq. (2) one finally obtains the path integral configuration partition function

$$Z_P = \left(\frac{m}{2\pi\hbar^2\beta_P} \right)^{P/2} \int \mathrm{d}q_1 \ldots \mathrm{d}q_P e^{-\beta_P \sum_{i=1}^{P} \left[V(q_i) + \frac{1}{2} m \omega_P^2 (q_i - q_{i+1})^2 \right]} \tag{4}$$

where cyclic boundary conditions are implied in the sum, $i + P \equiv i$.

[b] Note that we use the Trotter splitting $e^{A+B} \approx e^{A/2} e^B e^{A/2}$, which has a lower error than the asymmetric splitting $e^A e^B$.

Ignoring for a second the configuration-independent pre-factor, let us discuss the connection between Eq. (4) and the Feynmann path integral (1). Consider q_i to be a discrete sample from a continuous path, taken at $\tau_i = \beta\hbar i/P$. Then one can see the sum in the exponential as a discretization of a Riemann integral, and $(q_{i+1} - q_i)/(\tau_{i+1} - \tau_i)$ as a finite-difference approximation to $\dot{q}(\tau_i)$

$$\frac{\beta}{P}\sum_{i=1}^{P}\frac{(\tau_{i+1}-\tau_i)}{\beta\hbar/P}\left[V(q_i)+\frac{1}{2}m\frac{(q_i-q_{i+1})^2}{(\tau_{i+1}-\tau_i)^2}\right] \approx \frac{1}{\hbar}\int_0^{\beta\hbar} \mathrm{d}\tau \left[V(q(\tau))+\frac{1}{2}m\dot{q}(\tau)^2\right].$$

The multiple integrals over the q_j coordinates are the discrete equivalent of the path integral $\oint \mathcal{D}[q(\tau)]$. Even though most path integral simulations can be implemented and understood without reference to the formulation in terms of a functional integral, it is useful to keep this limit in mind, particularly when writing estimators for physical observables that are usually better behaved when they can be expressed as the discretized version of a corresponding path integral[14].

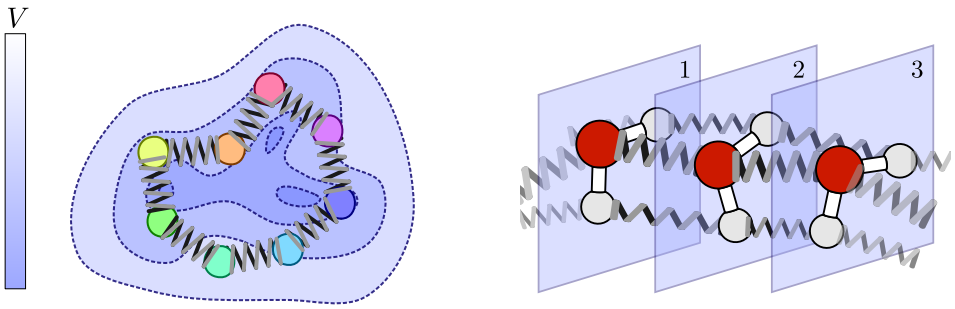


Figure 1: (Left panel) Cartoon representation of a classical ring polymer corresponding to the discretized path integral partition function (4). (Right panel) In a multi-atom setting, the ring polymer metaphor can be somewhat misleading. The path integral partition function is best seen as a sequence of imaginary-time slices: atoms within each replica interact with the physical potential, and the spring terms connect corresponding atoms in adjacent slices.

Eq. (4) corresponds precisely to the *classical* partition function of a cyclic polymer composed of P atoms, each of which is subject to the potential V and of a harmonic attractive interaction with its nearest neighbors. This isomorphism motivates the common practice of referring to the set of replicas for one atom as a "ring polymer" or a "necklace" and to each replica as a "bead". While this is a very suggestive metaphor, and can also be extended to give a pictorial representation of the Monte Carlo moves that are introduced to treat particle exchange effects[5], one has to keep in mind that in a real system it is better to regard the path integral partition function as describing a collection of "parallel universes", with atoms interacting with each other within each imaginary time slice and the kinetic part of the action corresponding to springs that connect each atom to its counterpart in the two adjacent time slices (see Figure (1)).

1.2 Estimators

Having defined an isomorphism between the quantum partition function of a system of distinguishable nuclei and the ring-polymer configurational partition function (4), one may proceed to evaluate experimental observables. To do so, one has to introduce appropriate *estimators*, functions of the coordinates of the ring polymer that correspond to physical quantities. The simplest case is that of the potential energy, or of any observable $\hat{A}(q)$ that depends solely on the atomic positions, such as a bond length, a radial distribution function, or the relative stability of two molecular configurations. The quantum mechanical expectation value can be written as $\langle A \rangle = \mathrm{Tr}\left[\hat{A}e^{-\beta\hat{H}}\right] / \mathrm{Tr}\, e^{-\beta\hat{H}}$. The operator $\hat{A}$ can be kept on the left, so that after the Trotter factorization and the splitting of the integral it appears close to the $\langle q_1|$ in Eq. (2), yielding a term $A(q_1)$. Since all the replicas are equivalent, one may as well average over the value of the observable computed on all the beads, so that the expectation value reads

$$\langle A \rangle_P = \frac{\int \mathrm{d}q_1 \dots \mathrm{d}q_P e^{-\beta_P \sum_{i=1}^{P}\left[V(q_i)+\frac{1}{2}m\omega_P^2(q_i-q_{i+1})^2\right]} \frac{1}{P}\sum_{i=1}^{P} A(q_i)}{\int \mathrm{d}q_1 \dots \mathrm{d}q_P e^{-\beta_P \sum_{i=1}^{P}\left[V(q_i)+\frac{1}{2}m\omega_P^2(q_i-q_{i+1})^2\right]}}. \tag{5}$$

This average can be computed easily by sampling the ring polymer configurations consistently with the ring polymer energy $\sum_{i=1}^{P}\left[V(q_i) + \frac{1}{2}m\omega_P^2(q_i - q_{i+1})^2\right]$ at the inverse temperature β_P, with Monte Carlo or (as it will be discussed in Section 2) molecular dynamics, and accumulating statistics for each replica. Note that this does *not* necessarily mean that averages will converge faster then if one was sampling just one replica, as in classical sampling: different path integral replicas are typically highly correlated with one another, which means that in most cases very little is gained by the average in (5).

While it is simple to write an estimator for observables that depend only on the positions, it is generally more complex to extract momentum-dependent quantities. A good example is that of the total energy, that contains both the position-dependent potential energy, but also a kinetic energy term. The simplest form of an estimator for the total energy of the system can be obtained recalling the thermodynamic relation between the partition function and the mean energy $\langle E \rangle = -Z^{-1}\partial Z/\partial\beta$. Once applied to Eq. (2), this reads

$$\langle E \rangle = \frac{\int \mathrm{d}q_1 \dots \mathrm{d}q_P e^{-\beta_P \sum_{i=1}^{P}\left[V(q_i)+\frac{1}{2}m\omega_P^2(q_i-q_{i+1})^2\right]} E^{\mathrm{TD}}(q_1, \dots q_P)}{\int \mathrm{d}q_1 \dots \mathrm{d}q_P e^{-\beta_P \sum_{i=1}^{P}\left[V(q_i)+\frac{1}{2}m\omega_P^2(q_i-q_{i+1})^2\right]}},$$

the ensemble average of the so-called *thermodynamic* (or *primitive*) energy estimator[c]

$$E^{\mathrm{TD}}(q_1, \dots q_P) = \frac{P}{2\beta} - \frac{1}{P}\sum_{i=1}^{P}\frac{1}{2}m\omega_P^2(q_i - q_{i+1})^2 + \frac{1}{P}\sum_{i=1}^{P}V(q_i).$$

One can immediately recognize the estimator for the potential energy, and infer that the thermodynamic estimator for the kinetic energy alone reads

$$T^{\mathrm{TD}}(q_1, \dots q_P) = \frac{P}{2\beta} - \frac{1}{P}\sum_{i=1}^{P}\frac{1}{2}m\omega_P^2(q_i - q_{i+1})^2. \tag{6}$$

[c] When doing the derivation keep in mind that $\omega_P = 1/\beta_P\hbar$ and that the constant scaling of Z_P also depends on β

Note that this estimator does not depend solely on the distribution of individual beads, but also on the cross-correlations between different replicas in the ring polymer. This is a general feature for non-local estimators that also contain a kinetic energy (or momentum) contribution.

Unfortunately, the thermodynamic kinetic energy estimator (6) is not very efficient, because its variance grows with the number of beads[15] as P/β^2. This means that computing the average to a given accuracy becomes more difficult as the number of replicas is increased, making the simulation even more computationally demanding.

Luckily, it is possible to exploit the virial theorem and do an integration by parts to derive the so-called centroid-virial kinetic energy estimator

$$T^{\mathrm{CV}}\left(q_1, \dots q_P\right) = \frac{1}{2\beta} + \frac{1}{2P}\sum_{i=1}^{P}\left(q_i - \bar{q}\right)\frac{\partial V}{\partial q_i}, \qquad \bar{q} = \frac{1}{P}\sum_{i=1}^{P} q_i, \tag{7}$$

that does not exhibit this pathological behavior of the fluctuations[15]. This case is a typical example of a recurring theme in path integral methods: one can write estimators that yield the same average value, but have very different statistical convergence properties. See for instance Ref.[16] for a discussion of efficient estimators for the heat capacity, Ref.[14] for an estimator of the distribution of particle momenta, and Refs.[17,18] for a comparison of different estimators for the isotope fractionation ratio – the relative propensity of the isotopes of the same element for different stable phases.

1.3 High-order Path Integrals

We have seen that Eq. (2) can be interpreted as a discretized form of a line integral over closed paths in configuration space. The error arising from using a finite P is effectively a discretization error, so one might wonder if it is possible to increase the order of convergence by employing a different summation rule. The crux is the error arising from splitting the exponential of the Hamiltonian, neglecting the commutator $\left[\hat{T}, \hat{V}\right]$. Hence, one can hope to increase the order of convergence by including extra terms that also depend on this commutator. A considerable amount of research has been devoted to this topic, in part also because of a connection with algorithms to propagate Hamiltonian dynamics in real time[19–22,16,23]. Among the many factorizations that have been proposed, one of the simplest and most successful is the Suzuki-Chin propagator

$$e^{-2\beta\hat{H}/P} = e^{-\frac{1}{3}\beta_P\hat{V}_{\mathrm{e}}}e^{-\beta_P\hat{T}}e^{-\frac{2}{3}\beta_P\hat{V}_{\mathrm{o}}}e^{-\beta_P\hat{T}}e^{-\frac{1}{3}\beta_P\hat{V}_{\mathrm{e}}} + \mathcal{O}\left(\beta_P^5\right), \tag{8}$$

in which one introduces two distinct modified potential energy operators that act on the odd and the even beads:

$$\hat{V}_{\mathrm{e}} = \hat{V} + \frac{\alpha}{6m\omega_P^2}\left|\frac{\partial V}{\partial q}\right|^2, \qquad \hat{V}_{\mathrm{o}} = \hat{V} + \frac{1-\alpha}{12m\omega_P^2}\left|\frac{\partial V}{\partial q}\right|^2.$$

The parameter $\alpha \in [0,1]$ can be tuned to optimize the prefactor for the discretization error. The square modulus of the force can typically be computed without additional effort, in particular when using molecular dynamics to sample the ring polymer partition function, as we will discuss in Section 2. However, the factorization (8) becomes inconvenient precisely when one wants to integrate the dynamics that arises from $\hat{V}_{\mathrm{e,o}}$: the derivative of

$|\partial V/\partial q|^2$ contains the second derivative of the potential, which often is considerably more computationally demanding than the force. In these cases, one can get around this difficulty by sampling the Trotter partition function, and using statistical reweighing to recover sampling consistent with the Suzuki-Chin modified potential[16,24–26]. This technique works very well for small clusters; however, the difference between the Trotter and the Suzuki-Chin potential energy is size-extensive, which means that it becomes progressively less efficient to include reweighing factors in the averages as the number of atoms is increased, somewhat limiting the applicability of this approach[25].

2 Path Integral Molecular Dynamics

As it has been already discussed above, the ring polymer partition function (2) only needs to depend explicitly on coordinates, and it can be sampled by Monte Carlo techniques. While Monte Carlo moves can be very effective in sampling phase space – and for instance they are used extensively in techniques to include particle exchange statistics[27,13] – they typically do not exploit the possibility of obtaining the inter-atomic forces with little overhead over the calculation of the potential energies. In many cases, particularly when one does not want to develop custom-tailored Monte Carlo moves for the specific system at hand, it can be much more effective to sample the Boltzmann distribution by integrating in time Hamilton's equations

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{p}{m}, \qquad \dot{p} = -\frac{\partial H}{\partial q} = -\frac{\partial V}{\partial q}. \tag{9}$$

In practice, it is easy to see that the constant pre-factor in Eq. (2) corresponds to a Gaussian integral over a set of auxiliary variables that can be taken to be the conjugate momenta to q_i's, so that one can equivalently write the ring polymer partition function as

$$Z_P = \frac{1}{(2\pi\hbar)^P} \int \mathrm{d}p_1 \ldots \mathrm{d}p_P \int \mathrm{d}q_1 \ldots \mathrm{d}q_P e^{-\beta_P \sum_{i=1}^{P} \left[V(q_i) + \frac{p_i^2}{2m} + \frac{1}{2} m\omega_P^2 (q_i - q_{i+1})^2 \right]}. \tag{10}$$

Note that the momenta p_i are exclusively sampling devices, *and are in no way related to the physical momentum*. Among other things, this means that one could change the inertial mass in the $p_i^2/2m$ term to be something different from the physical mass, with no other effect than changing the partition function by an immaterial, temperature-independent scaling.

2.1 Implementation: Normal Modes Propagator

While the underlying idea behind path integral molecular dynamics (PIMD) is very simple, there are some technical aspects that should be considered in order to obtain an efficient implementation. Start by splitting the path integral Hamiltonian into a "free-particle" component, and one that depends on the physical potential:

$$H_P = H_0 + V_P = \sum_{i=1}^{P} \left[\frac{p_i^2}{2m} + \frac{1}{2} m\omega_P^2 \left(q_i - q_{i+1} \right)^2 \right] + \sum_{i=1}^{P} V\left(q_i \right).$$

The free ring-polymer Hamiltonian H_0 is just a multi-dimensional harmonic oscillator, and therefore it can be diagonalized exactly by the unitary transformation

$$\tilde{p}_j = \sum_i p_i C_{ij}, \quad \tilde{q}_j = \sum_i q_i C_{ij}, \quad C_{ij} = \sqrt{\frac{2}{P}} \cdot \begin{cases} 1/\sqrt{2} & j = 0 \\ \cos 2\pi ij/P & j < P/2 \\ (-1)^i/\sqrt{2} & j = P/2 \\ \sin 2\pi ij/P & j > P/2 \end{cases}.$$

$\tilde{q}_j$ and $\tilde{p}_j$ are the position and momentum of the j-th free-particle normal mode, which has an associated frequency

$$\omega_j = 2\omega_P \sin j\pi/P. \tag{11}$$

Considering that for a physical system that contains a normal mode of frequency $\omega_{\max}$ one has to use at least $2\beta\hbar\omega_{\max}$ replicas, a converged PIMD calculation will involve normal-mode vibrational frequencies larger than $4\omega_{\max}$. This implies that in principle one should use a much smaller time step to integrate the equations of motion for PIMD compared to conventional classical molecular dynamics, adding further computational burden to the technique.

Luckily, one can exploit the possibility of diagonalizing the free ring-polymer Hamiltonian to avoid reducing the time step. This can be achieved by doing a so-called staging transformation[28], which will not be discussed here, or by performing the integration in the normal modes basis. In the latter case, one can for instance manipulate the inertial mass of the conjugate momenta in the normal modes representation, so as to artificially slow down the dynamics of the fast ring-polymer normal modes without changing the sampling properties. In most cases, however, this is not necessary: one can perform a multiple time step procedure[29] to integrate the normal modes analytically based on the evolution of a free ring polymer and then include the physical potential using a velocity Verlet algorithm[30].

This is based on the symmetric Trotter splitting of the Liouville operator for H_P into the part related to H_0 and that related to the physical potential $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_V$. Considering the evolution of the ring polymer over a time step Δt

$$e^{-\Delta t \mathcal{L}} \approx e^{-\Delta t \mathcal{L}_V/2} e^{-\Delta t \mathcal{L}_0} e^{-\Delta t \mathcal{L}_V/2}.$$

This corresponds to the following recipe for the evolution across a time step:

$$\begin{aligned} p_i &\leftarrow p_i - \frac{\Delta t}{2} \frac{\partial V(q_i)}{\partial q_i} \\ \tilde{p}_j &\leftarrow \sum_i p_i C_{ij} \qquad \tilde{q}_j \leftarrow \sum_i q_i C_{ij} \\ \begin{pmatrix} \tilde{p}_j \\ \tilde{q}_j \end{pmatrix} &\leftarrow \begin{pmatrix} \cos\omega_j \Delta t & -m\omega_j \sin\omega_j \Delta t \\ [1/m\omega_j] \sin\omega_j \Delta t & \cos\omega_j \Delta t \end{pmatrix} \begin{pmatrix} \tilde{p}_j \\ \tilde{q}_j \end{pmatrix} \\ p_i &\leftarrow \sum_j C_{ij} \tilde{p}_j \qquad q_i \leftarrow \sum_j C_{ij} \tilde{q}_j \\ p_i &\leftarrow p_i - \frac{\Delta t}{2} \frac{\partial V(q_i)}{\partial q_i}. \end{aligned} \tag{12}$$

In practice, one could reduce the number of normal modes transformations (that can be performed by fast Fourier transform, and therefore are not dramatically demanding anyway) by keeping the momenta in the normal modes representation, and transforming the physical forces in the normal modes basis. See for instance Ref.[31] for a discussion of Eq. (12) in a many-atoms context.

It is also worth mentioning that whenever the inter-atomic potential can be split in a part that varies on a short length scale (such as intra-molecular bends and stretches) and a long-range scale (such as electrostatics), the normal-modes representation can also be exploited within a *ring-polymer contraction* scheme[32,33], that reduces the cost of the simulation by computing the long-range part of the potential on a reduced number of replicas.

2.2 Efficient Stochastic Thermostatting

Molecular dynamics generates trajectories that are consistent with a Boltzmann distribution. However, they do not yield ergodic sampling, because they conserve the total energy and so a single MD simulation cannot account for the thermal fluctuations that are characteristic of a constant-temperature ensemble, and that are needed to converge averages based on the partition function (10). This fact has been recognized for a long time, and led to the development of modified dynamical equations that describe the heat exchange with a reservoir, and generate ergodic canonical trajectories[34–40]. In principle, any of these techniques could be applied to PIMD, which is just classical molecular dynamics in an extended phase space. Perhaps, the simplest technique to be used in conjunction with the normal-modes propagation in Eq. (12) is one that applies Langevin dynamics in the normal modes representation, so that the friction can be tuned to match the critical-damping value for the free ring-polymer normal modes. We introduce a Langevin term in the equation of motion for the momenta, corresponding to a further term $\mathcal{L}_\gamma$ in the Liouvillian, that can therefore be split as

$$e^{-\Delta t\mathcal{L}} \approx e^{-\Delta t\mathcal{L}_\gamma/2} e^{-\Delta t\mathcal{L}_V/2} e^{-\Delta t\mathcal{L}_0} e^{-\Delta t\mathcal{L}_V/2} e^{-\Delta t\mathcal{L}_\gamma/2}.$$

This splitting corresponds to applying the following step

$$\dot{\tilde{p}}_j \leftarrow e^{-\gamma_j \Delta t/2}\tilde{p}_j + \sqrt{\left(1-e^{-\gamma_j\Delta t}\right) m/\beta_n}\xi_j$$

twice per time step, immediately before and immediately after the Hamiltonian propagator (12). ξ_j's are uncorrelated Gaussian random numbers, such that $\langle \xi_j(t)\,\xi_{j'}(t')\rangle = \delta_{jj'}\delta(t-t')$. Note that in the most naive implementation this step requires also to transform forth and back to normal modes representation, a problem that is easily circumvented by propagating the momenta in the normal modes form. The friction can be taken to be $\gamma_j = \lambda\omega_j$, where ω_js are the free ring polymer normal mode frequencies, and λ is a scaling that can be set to one to have critical damping of harmonic oscillations, or to a smaller value when one needs a more gentle thermostatting. The centroid mode (that has zero frequency in the free particle limit) can be treated separately, using either a white-noise Langevin thermostat or stochastic velocity rescaling[39,31].

2.3 Approximate Quantum Dynamics

Even though, strictly speaking, path integral molecular dynamics is just a sampling technique to obtain static, equilibrium averages, it can be used as the basis to compute ap-

proximate time correlation functions that include some of the nuclear quantum effects on dynamical properties. Examples of time-dependent properties that can be estimated in this way include diffusion coefficients, reaction rates, vibrational spectra. The two main techniques are centroid molecular dynamics (which amounts at performing microcanonical molecular dynamics on the centroid potential of mean force[11,41]) and ring-polymer molecular dynamics (which corresponds to molecular dynamics on the ring polymer potential energy surface, using the physical masses in the definition of the Hamiltonian[42,43,12]). The use of both these techniques can be partially justified based on how they can capture quantum mechanical behavior in some limits, or as approximations to more rigorous formulations of quantum dynamics[44–48]. They typically behave similarly for dynamical properties that evolve on a long time scale, but exhibit evident artifacts for short-time dynamics, e.g. for the stretching modes in IR spectra[49]. Interestingly, using physical masses but thermostatting the internal modes of the ring polymer – an approach that is half-way between the two techniques – eliminates the most apparent artifacts[50,51], providing a reliable albeit approximate method to probe the impact of quantum nuclei on vibrational properties.

3 Accelerating Convergence with Colored Noise

An alternative approach to accelerate the convergence of physical observables in path integral molecular dynamics combines a simulation with a small number of replicas with correlated-noise Langevin dynamics[52,53]. The idea is that a non-equilibrium Langevin dynamics generates frequency-dependent fluctuations, that can be used to mimic the effect of zero-point energy on a multi-dimensional harmonic oscillator, without the need of knowing the normal modes frequencies or eigenvectors. By combining colored noise with path integral molecular dynamics one can obtain a simulation protocol that, in the harmonic limit, yields exact results for any number of beads. As the number of replicas is increased, this protocol converges by construction to PIMD, and so the method can be made as accurate as one wishes, and in most cases has smaller errors than plain PIMD for any number of replicas.

3.1 A Brief Introduction to Colored Noise

Stochastic differential equations (SDEs) combine a random/noisy term together with deterministic prescriptions for the time evolution of a system[54]. The prototypical example for the application of SDEs to a physical problem is the Langevin equation, that was originally introduced as a model for Brownian motion[55], and has been used extensively to describe the coupling between an open system and an external bath[56]. Langevin dynamics introduce a friction term $-\gamma p$ and a Gaussian random force ξ on top of Hamiltonian dynamics

$$\begin{aligned} \dot{q} =& p/m \\ \dot{p} =& -V'(q) - \gamma p + \sqrt{2m\gamma/\beta}\xi(t). \end{aligned} \tag{13}$$

The balance between the noisy force (which is uncorrelated in time, $\langle \xi(t)\,\xi(t')\rangle = \delta(t-t')$) and the friction guarantees canonical sampling at inverse temperature β.

If the problem of integrating out the degrees of freedom associated with the bath is considered carefully[57,58], one will find that in the general case it is not possible to represent

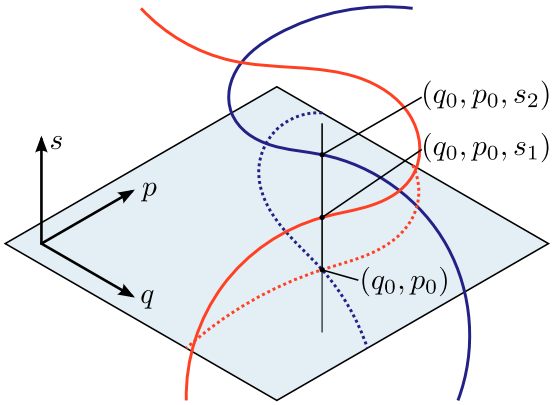


Figure 2: Two Markovian trajectories in the extended (q, p, s) space will appear as history-dependent, non-Markovian trajectories when projected into the (q, p) subspace.

the effect of the bath by equations as (13), but that the friction and the noisy force have to be associated to a finite memory kernel, leading to the non-Markovian dynamics

$$\begin{aligned} \dot{q} &= p/m \\ \dot{p} &= -V'(q) - \int_{-\infty}^{t} K(t-s)p(s)\mathrm{d}s + \sqrt{2m/\beta}\zeta(t). \end{aligned} \tag{14}$$

The autocorrelation of the noisy force $H(t) = \langle \zeta(t)\zeta(0) \rangle$ must be related to the memory of the friction $K(t)$ by a fluctuation-dissipation relation $H(t) = K(t)$ to guarantee constant-temperature sampling. Besides giving the possibility of modeling the statistical mechanics and the dynamical behavior of an open system[59,60], a generalized Langevin equation of this form gives a great deal of freedom for manipulating the sampling properties of a molecular dynamics simulation by changing the form of the memory kernels $K(t)$ and $H(t)$. However, integrating an equation of motion of this form is not very practical, since one would need to generate Gaussian random numbers with prescribed correlation[61], to store the past trajectory of $p(t)$ and to perform an integral over the past history to compute the friction term.

Fortunately, one can reverse the reasoning that leads to Eqs. (14) starting from Markovian equations for the combined system/bath ensemble to show that a non-Markovian dynamics for (q, p) can be obtained by supplementing the physical variables with a vector $\mathbf{s}$ of n_s fictitious momenta[62–64], and integrating the Markovian equations

$$\begin{aligned} \dot{q} =& p/m \\ \begin{pmatrix} \dot{p} \\ \dot{\mathbf{s}} \end{pmatrix} =& \begin{pmatrix} -V'(q) \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} a_{pp} & \mathbf{a}_p^T \\ \bar{\mathbf{a}}_p & \mathbf{A} \end{pmatrix} \begin{pmatrix} p \\ \mathbf{s} \end{pmatrix} + \sqrt{m} \begin{pmatrix} b_{pp} & \mathbf{b}_p^T \\ \bar{\mathbf{b}}_p & \mathbf{B} \end{pmatrix} \begin{pmatrix} \xi \\ \boldsymbol{\xi} \end{pmatrix}, \end{aligned} \tag{15}$$

in the extended phase space comprising $(q, p, \mathbf{s})$. The additional degrees of freedom encode the memory of the system relative to the (q, p) physical variables - e.g. two trajectories

starting at the same (q,p) position (that would necessarily evolve in the same way in a Markovian context) can now lead to different time evolution, if they correspond to different values of the $\mathbf{s}$ momenta (Figure 2). Let $\mathbf{A}_p$ and $\mathbf{B}_p$ be the full $(n_s+1)\times(n_s+1)$ matrices in Eq. (15). The fluctuation-dissipation relation that guarantees canonical sampling is then $\mathbf{A}_p + \mathbf{A}_p^T = \beta^{-1}\mathbf{B}_p\mathbf{B}_p^T$. Eq. (15) corresponds to a non-Markovian dynamics (14) with $K(t) = 2a_{pp}\delta(t) - \mathbf{a}_p^T e^{-|t|\mathbf{A}}\bar{\mathbf{a}}_p$.

Formulating the Generalized Langevin Equation (GLE) in this Markovian form is not only convenient for the sake of integrating the dynamics on a computer. If one considers a harmonic model, for which the force $V'(q) = m\omega^2 q$ is linear, the whole set of dynamical equations for $(q,p,\mathbf{s})$ constitutes a linear, Markovian stochastic differential equation (an Ornstein-Uhlenbeck process[54]) that can be solved analytically. The opportunity of computing exactly and inexpensively the response of a normal mode of frequency ω to a given set of GLE parameters $\mathbf{A}_p$ and $\mathbf{B}_p$ makes it possible to optimize iteratively the parameters to fulfill the desired sampling properties (e.g. ergodicity, small disturbance on the dynamics, etc.) as a function of the normal mode frequency.

A crucial feature of Eqs. (15) that makes this approach very useful is that, when applying identical equations with independent random terms to a set of arbitrarily-coupled degrees of freedom, the dynamical behavior is invariant to a unitary transformation of the coordinates to which the stochastic dynamics is applied. This property, that is a consequence of the linear nature of the SDE and of the Gaussian statistics of $\boldsymbol{\xi}$, means that one can apply the GLEs to the Cartesian coordinates of a set of atoms and obtain the same response as if the GLEs had been applied to the normal modes of the system. Hence, one can optimize $\mathbf{A}_p$ and $\mathbf{B}_p$ based on a simple, one-dimensional harmonic oscillator model, but the predictions will be verified in a realistic physical model without the need of knowing explicitly the normal modes basis. This idea has been applied to obtain efficient sampling of many frequencies at the same time, and to avoid disrupting adiabatic decoupling in Car-Parrinello molecular dynamics[40,65], but also to thermostat efficiently PIMD[31] and to stabilize resonances in multiple time step integration[66].

3.2 Quantum thermostats

Generalized Langevin dynamics that satisfy the fluctuation-dissipation relation can be used to alter the dynamical properties of a trajectory, and also to make sampling more efficient in PIMD[31]. By releasing the constraint $\mathbf{A}_p + \mathbf{A}_p^T = \beta^{-1}\mathbf{B}_p\mathbf{B}_p^T$, one is effectively in a non-equilibrium scenario, that can be understood in terms of simultaneous coupling to multiple heat baths at different temperature. The practical effect of this simulation protocol is that a steady state will be reached, in which normal modes of different frequency will have fluctuations that are consistent with different values of the temperature. This peculiar non-equilibrium GLE is particularly useful to model inexpensively the quantum nature of atomic nuclei.

To see how, consider a one-dimensional harmonic oscillator of frequency ω, sampled canonically at inverse temperature $\beta = 1/k_B T$. The phase-space distribution of position and momentum are Gaussian distributions $\rho(p) \propto \exp -p^2/2\sigma_p^2$ and $\rho(q) \propto \exp -q^2/2\sigma_q^2$, irrespective of the underlying classical or quantum description. However, the classical and quantum distributions differ because of the magnitude of fluctuations: for a classical oscillator $\sigma_p^2 = m/\beta$ and $\sigma_q^2 = 1/\beta m\omega^2$, while for a quantum oscilla-

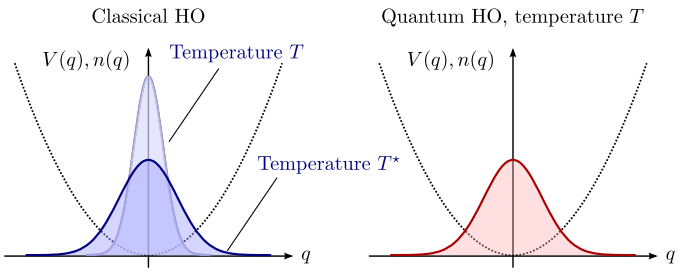


Figure 3: The phase-space distribution of positions for a harmonic oscillator at finite temperature T is a Gaussian, both in classical and quantum mechanical treatments. The only difference between the two cases is the amplitude of the distribution. Quantum fluctuations can be mimicked in a classical context by increasing the temperature $T^\star$, that however will be frequency dependent.

tor $\sigma_p^2 = m\frac{\hbar\omega}{2}\coth\frac{\beta\hbar\omega}{2}$ and $\sigma_q^2 = \frac{\hbar}{2m\omega}\coth\frac{\beta\hbar\omega}{2}$. A classical oscillator would exhibit the very same behavior as the quantum mechanical one if it were modeled at an effective temperature

$$T^\star(\omega) = \frac{\langle p^2\rangle}{mk_B} = \frac{m\langle q^2\rangle}{k_B} = \frac{\hbar\omega}{2k_B}\coth\frac{\hbar\omega}{2k_BT}. \tag{16}$$

such that the fluctuations match those predicted by quantum mechanics at temperature T (see Figure 3).

The problem with this idea is that the effective temperature $T^\star(\omega)$ depends on the oscillator frequency as well as on the physical target temperature. In the case of a multi-dimensional oscillator (that can be taken to be a decent model for a solid at low temperature) one would need to associate a different temperature to each normal mode. If one wanted to do so using conventional thermostatting, one would need to know the normal modes frequencies and phonon displacement patterns, and apply tailored white-noise thermostats at different temperatures working in the normal modes representation. The advantage of a non-equilibrium GLE formulation, instead, is that the dynamical response of the system determines the fluctuations along different directions automatically, without the need of knowing the normal modes of the system being studied. If one can fit a set of $\mathbf{A}_p$ and $\mathbf{B}_p$ parameters that enforce the frequency-dependent temperature (16), for any frequency within a range that encompasses the vibrational modes relevant for the system at hand, then it suffices to apply the same GLE to each Cartesian degree of freedom. The quantum $T^\star(\omega)$ curve is then enforced automatically, giving quantum fluctuations at the cost of conventional molecular dynamics.

This "quantum thermostat" (QT) idea[67,65] works surprisingly well also for strongly anharmonic potentials, and the main limitation when applying it to real systems does not depend much on failure to describe strongly anharmonic behaviour. One can see that large deviations from quantum behavior are caused by zero-point energy leakage, i.e. the fact that due to weak anharmonic couplings energy flows from the fast normal modes, that are thermostatted at high $T^\star$ to account for the large zero-point energy, to slow normal modes that are nearly classical[65]. This energy flow was not accounted for when designing $T^\star(\omega)$, and so there will be a (significant) deviation between the desired quasi-harmonic quan-

tum fluctuations and the actual fluctuations. This problem is common to semi-classical methods to treat quantum dynamics[68], and has been also recognized in other stochastic approaches to obtain approximate quantum effects[69,70]. Rather than trying to remedy this problem by exploiting information on the anharmonic couplings – which would be an ad-hoc, non-transferable solution, requiring in-depth knowledge of the system – one can control zero-point energy leakage by exploiting the tunability of the GLE thermostats, enforcing a strong coupling across the whole frequency range so as to counterbalance effectively the zero-point energy leakage. This approach improves significantly the performance of the quantum thermostat when applied to anharmonic problems[65,71], and made it possible to describe semi-quantitatively the role of NQEs in several real applications[72,73].

3.3 Combining Generalized Langevin Equations and PIMD

The approximations behind the quantum thermostat and related semi-classical methods are basically uncontrolled, and very hard to gauge unless it is possible to perform a harmonic analysis. Therefore, the quantum thermostat can be regarded as an inexpensive technique to assess qualitatively the importance of NQEs, but it is not recommend if one wants to infer quantitative conclusions. One could then imagine to combine colored-noise and path integral molecular dynamics: the former is only exact in the harmonic limit, while the latter converges systematically but often requires a large number of replicas and is therefore computationally demanding.

The crux is designing a GLE thermostat that enforces exact quantum fluctuations in the harmonic limit *for any number of replicas*, even in cases where PIMD alone would be far from converged. Such a PI+GLE method inherits from the quantum thermostat the property of being exact for harmonic problems, and naturally converge to (Boltzmann-sampled) PIMD when the number of beads is large enough to have a converged result in the absence of correlated noise.

In order to work out the properties of the GLE that would achieve this goal, one can proceed in the same way as with the quantum thermostat, only considering that now, in the presence of a harmonic potential of frequency ω, the ring-polymer normal mode frequencies will be changed from the free-particle value (11) and become $\omega_j = \sqrt{\omega^2 + 4P^2 \sin^2 j\pi/P}$. These are the frequencies that will be picked up by the colored-noise dynamics. Introducing a frequency-dependent *configurational* temperature $T^*(\omega) = \left\langle q^2 \right\rangle(\omega) m\omega^2/k_B$ (momentum fluctuations are not important per se in a PIMD framework), one gets the requirement for having quantum fluctuations of the beads to be

$$\begin{aligned} \frac{m\omega^2}{k_B T}\left\langle q^2 \right\rangle &= \frac{m\omega^2}{Pk_B T}\sum_i \left\langle q_i^2 \right\rangle = \frac{m\omega^2}{Pk_B T}\sum_j \left\langle \tilde{q}_j^2 \right\rangle \\ &= \frac{1}{P}\sum_j \frac{T^*\left(\omega_j\right)/T}{\omega_j^2/\omega^2} = \frac{\hbar\omega}{2k_B T}\coth\frac{\hbar\omega}{2k_B T}. \end{aligned} \tag{17}$$

Since ω_j depends on the physical frequency ω, Eq. (17) must be seen as a functional equation that defines the $T^*\left(\omega\right)$ curve. As shown in Ref.[52], Eq. (17) can be conveniently written as a function of the a-dimensional parameter $x = \beta\hbar\omega/2$, that expresses qualitatively how much an oscillator deviates from a classical behavior, and can be solved numerically with

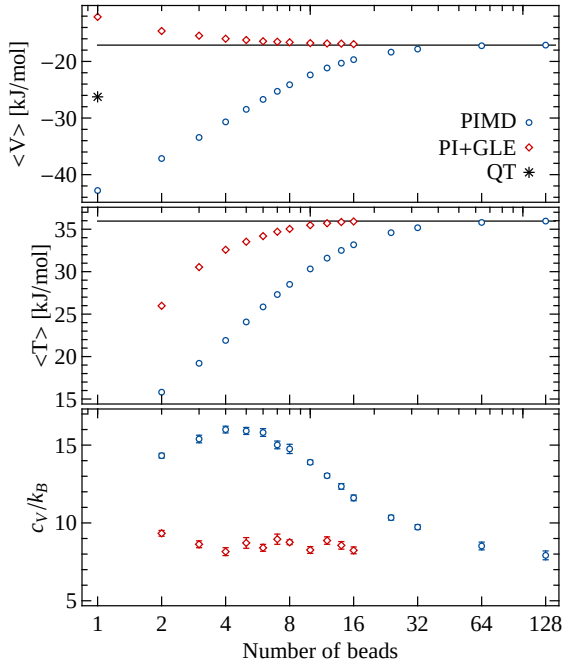


Figure 4: The average value of the potential energy, the virial kinetic energy and the constant-volume heat capacity for a simulation of a flexible water model[74] at T=298 K, plotted as a function of the number of beads. The results obtained with conventional PIMD and PI+GLE are compared, and the value of V obtained with the original quantum thermostat (QT) is also reported. (Adapted from Ref.[52])

an iterative technique. As the number of path integral replicas is increased, the curve remains constant up to a larger value of x, that corresponds to the fact that PI+GLE behaves as conventional PIMD with Boltzmann sampling of the ring-polymer Hamiltonian for oscillators with larger and larger frequency. This implies that PI+GLE is bound to converge to the exact quantum averages, just because in the large P limit it converges to PIMD.

Figure 4 shows the convergence with number of beads of potential and kinetic energy for a quantum simulation of an empirical water model[74] at room temperature, comparing plain PIMD and PI+GLE. Colored noise accelerates dramatically the convergence of observables to the quantum expectation values, and the possibility of converging results systematically makes it possible to assess the error. A careful examination of Figure 4 shows that the mean kinetic energy $\langle T \rangle$ converges somewhat more slowly than $\langle V \rangle$.

This is due to a specific shortcoming of PI+GLE, that becomes clear when one considers the expression for the centroid-virial kinetic energy estimator (7) in a one-dimensional harmonic potential, that reads:

$$\begin{aligned} \langle T \rangle =& \frac{1}{2\beta} + \frac{1}{2P}\omega^2 \sum_{i=0}^{P-1} \left\langle q_i^2 \right\rangle - \frac{1}{2}\omega^2 \left\langle \bar{q}^2 \right\rangle \\ =& \langle V \rangle + \frac{1}{2\beta} - \frac{1}{2}\omega^2 \left\langle \bar{q}^2 \right\rangle . \end{aligned} \tag{18}$$

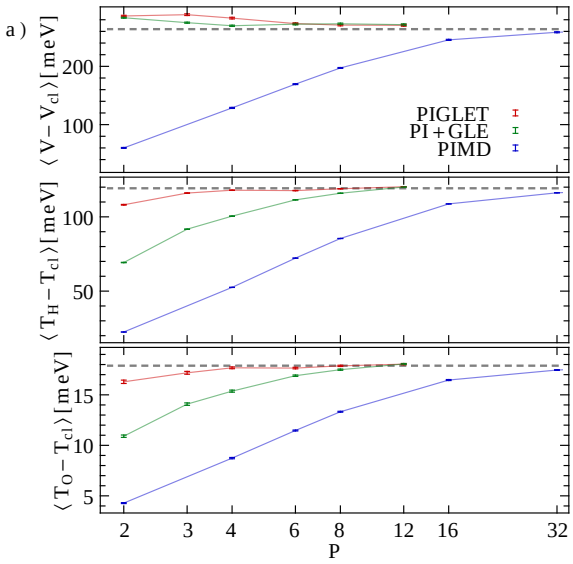


Figure 5: The quantum contribution to the potential energy, and to the kinetic energy of hydrogen and oxygen atoms as computed by the centroid virial estimator for a simulation of a flexible water model[74] at T=298 K, plotted as a function of the number of beads. Note the much accelerated convergence rate of the kinetic energy when using PIGLET compared to PI+GLE. (Adapted from Ref.[53])

The quantum mechanical expectation values for potential and kinetic energy of a harmonic oscillator of frequency ω read

$$\langle V \rangle = \langle T \rangle = \frac{\hbar\omega}{4} \coth \frac{\beta\hbar\omega}{2}, \tag{19}$$

so it is not sufficient that the fluctuations of q are consistent with $\langle V \rangle = \frac{\hbar\omega}{4} \coth \frac{\beta\hbar\omega}{2}$, but it is also necessary to make sure that the fluctuations of the centroid satisfy $\frac{1}{2}\omega^2 \langle \bar{q}^2 \rangle = \frac{1}{2\beta}$. This observation is a sign of a general limitation of the PI+GLE scheme, that enforces only quantum fluctuations for the "marginal" distribution of the beads, which is sufficient to guarantee accelerated convergence of any observable that depends only on q but does not necessarily help converging more complex estimators that also depend on the correlations between different beads.

Fortunately, it is relatively easy to extend the PI+GLE idea to include further correlations. When PIMD is propagated in the normal modes representation as discussed in Section 2 one can apply different thermostats to the various normal modes, and tune them separately to warrant faster convergence of multiple estimators simultaneously. This idea is used for instance in the PIGLET scheme[53] to converge simultaneously structural observables and the centroid-virial kinetic energy estimator, obtaining a further improvement over the PI+GLE scheme, as shown in Figure 5.

4 Conclusive Remarks

Atomistic simulations of ever-increasing accuracy are making it possible to model molecules and materials with predictive accuracy. One of the fundamental physical effects that is still ignored in many simulations involves the deviations from classical behavior of the atomic nuclei, that is very pronounced, at room temperature and below, for the lightest elements – hydrogen in particular. The importance of nuclear quantum effects has been demonstrated very clearly, for instance, by pioneering path integral molecular dynamics studies of water and charged water defects[75,10,76,77], and it is essential when it comes to compare with experiments – such as deep inelastic neutron scattering[2,78,14,72,79], or isotope fractionation measurements[80–83,17,84] – for which the quantum mechanical nature of nuclei is the very origin of the observed signal. Nevertheless in many modern simulations of water and H-containing compounds nuclear quantum effects are still ignored. This is at times justified: in neat water quantum effects along different molecular directions cancel out almost perfectly[74,85,86], and so overall NQEs on many thermodynamic quantities are small. However, simulations that aim at probing the dissociation of covalent O-H bonds[87,88], studying self-ionization or re-combination of hydronium and hydroxide[89], and more in general quantitatively characterizing the behavior of hydrogen bonds in unusual environments (high pressure, confinement, interfaces, solvation shells of ions, ...) should always at the very least assess whether NQEs do or do not play a major role. In many cases, the only excuse for not including NQEs is the large computational overhead involved in path integral simulations, combined with the technical complexity of the method and the fact that few up-to-date, well-maintained implementations exist in mainstream atomistic simulation packages. In this lecture we have provided a concise but complete introduction to the theory and the practical implementation of atomistic modeling of the quantum nature of atomic nuclei by path integral molecular dynamics. When one looks beyond its cumbersome notation, PIMD is just classical dynamics in an extended phase space, which means that all the tricks that are used in classical simulations (including e.g. thermostats for sampling the canonical ensemble[9], or barostats for simulations at constant-pressure[90]) can be applied transparently. What is more, simulating NQEs does not need to involve an increase of several orders of magnitude in computational effort. High-order factorizations of the quantum partition function lead to accelerated convergence in some circumstances, and correlated-noise thermostats give an accurate estimate of NQEs in water at room temperature with as little as six path integral beads.

The implementation burden that has been traditionally associated with path integral simulations is also being reduced by the introduction of modularly designed packages that deal with the PIMD aspects of the simulation while delegating to external codes the evaluation of energy and forces. An example of this concept is given by i-PI[88], an open-source[d] Python package that is already interfaced with the development versions of CP2K[91,92], LAMMPS[93] and FHI-AIMS[94]. When one also considers that the evaluation of energy and forces over multiple PI replicas involves a trivial layer of concurrency, that is ideally adapted to massively parallel high-performance computing, one can conclude that time is ripe for making the modeling of nuclear quantum effects in hydrogen-containing liquids and solids routine.

[d] `http://epfl-cosmo.github.io/gle4md/index.html?page=ipi`

References

1. M. Born and R. Oppenheimer, *Zur Quantentheorie der Molekeln*, Ann. Phys., **389**, 457–484, 1927.
2. C. Andreani, D. Colognesi, J. Mayers, G. F. Reiter, and R. Senesi, *Measurement of momentum distribution of light atoms and molecules in condensed matter systems using inelastic neutron scattering*, Adv. Phys., **54**, 377–469, 2005.
3. Gregory A Voth, *Calculation of equilibrium averages with Feynman-Hibbs effective classical potentials and similar variational approximations*, Phys. Rev. A, **44**, 5302–5305, 1991.
4. R P Feynman and A R Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1964.
5. David Chandler and Peter G Wolynes, *Exploiting the isomorphism between quantum theory and classical statistical mechanics of polyatomic fluids*, J. Chem. Phys., **74**, 4078–4095, 1981.
6. M Parrinello and A Rahman, *Study of an F center in molten KCl*, J. Chem. Phys., **80**, 860, 1984.
7. B J Berne and D Thirumalai, *On the Simulation of Quantum Systems: Path Integral Methods*, Annu. Rev. Phys. Chem., **37**, 401–424, 1986.
8. Dominik Marx and Michele Parrinello, *Ab initio path integral molecular dynamics: Basic ideas*, J. Chem. Phys., **104**, 4077, 1996.
9. Mark E Tuckerman, Dominik Marx, Michael L Klein, and Michele Parrinello, *Efficient and general algorithms for path integral Car-Parrinello molecular dynamics*, J. Chem. Phys., **104**, 5579–5588, 1996.
10. D Marx, M E Tuckerman, J Hutter, and M Parrinello, *The nature of the hydrated excess proton in water*, Nature, **397**, 601–604, 1999.
11. Jianshu Cao and Gregory A Voth, *A new perspective on quantum time correlation functions*, J. Chem. Phys., **99**, 10070–10073, 1993.
12. Scott Habershon, David E Manolopoulos, Thomas E Markland, and Thomas F Miller, *Ring-polymer molecular dynamics: quantum effects in chemical dynamics from classical trajectories in an extended phase space.*, Annu. Rev. Phys. Chem., **64**, 387–413, 2013.
13. D M Ceperley, *Path integrals in the theory of condensed helium*, Rev. Mod. Phys., **67**, 279–355, 1995.
14. Lin Lin, Joseph A Morrone, Roberto Car, and Michele Parrinello, *Displaced Path Integral Formulation for the Momentum Distribution of Quantum Particles*, Phys. Rev. Lett., **105**, 110602, 2010.
15. M. F. Herman, E J Bruskin, and B J Berne, *On path integral Monte Carlo simulations*, J. Chem. Phys., **76**, 5150, 1982.
16. Takeshi M Yamamoto, *Path-integral virial estimator based on the scaling of fluctuation coordinates: Application to quantum clusters with fourth-order propagators*, J. Chem. Phys., **123**, 104101, 2005.
17. Michele Ceriotti and Thomas E Markland, *Efficient methods and practical guidelines for simulating isotope effects.*, J. Chem. Phys., **138**, 014112, 2013.
18. Bingqing Cheng and Michele Ceriotti, *Direct path integral estimators for isotope fractionation ratios.*, J. Chem. Phys., **141**, 244112, 2014.

19. M Takahashi and M Imada, *Monte Carlo calculation of quantum systems. II. Higher order correction*, J. Phys. Soc. Japan, **53**, 3765–3769, 1984.
20. X P Li and J Q Broughton, *High-order correction to the Trotter expansion for use in computer simulation*, J. Chem. Phys., **86**, 5094, 1987.
21. M Suzuki, *Hybrid exponential product formulas for unbounded operators with possible applications to Monte Carlo simulations*, Phys. Lett. A, **201**, 425–428, 1995.
22. Seogjoo Soonmin Jang and Gregory A Voth, *Applications of higher order composite factorization schemes in imaginary time path integral simulations*, J. Chem. Phys., **115**, 7832–7842, 2001.
23. Marcin Buchowiecki and Jirí Vanícek, *Monte Carlo evaluation of the equilibrium isotope effects using the Takahashi-Imada factorization of the Feynman path integral*, Chem. Phys. Lett., **588**, 11–16, 2013.
24. Alejandro Pérez and Mark E Tuckerman, *Improving the convergence of closed and open path integral molecular dynamics via higher order Trotter factorization schemes.*, J. Chem. Phys., **135**, 064104, 2011.
25. Michele Ceriotti, Guy a. R. Brain, Oliver Riordan, and David E. Manolopoulos, *The inefficiency of re-weighted sampling and the curse of system size in high-order path integration*, Proc. R. Soc. A Math. Phys. Eng. Sci., **468**, 2–17, 2011.
26. Ondrej Marsalek, Pei-Yang Chen, Romain Dupuis, Magali Benoit, Merlin Méheut, Zlatko Bačić, and Mark E. Tuckerman, *Efficient Calculation of Free Energy Differences Associated with Isotopic Substitution Using Path-Integral Molecular Dynamics*, J. Chem. Theory Comput., **10**, 1440–1453, 2014.
27. DM M Ceperley and EL L Pollock, *Path-integral computation of the low-temperature properties of liquid 4He*, Phys. Rev. Lett., **56**, 351–354, 1986.
28. Michiel Sprik, Michael Klein, and David Chandler, *Staging: A sampling technique for the Monte Carlo evaluation of path integrals*, Phys. Rev. B, **31**, 4234–4244, 1985.
29. M. Tuckerman, B. J. Berne, and G. J. Martyna, *Reversible multiple time scale molecular dynamics*, J. Chem. Phys., **97**, 1990, 1992.
30. Loup Verlet, *Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules*, Phys. Rev., **159**, 98–103, 1967.
31. Michele Ceriotti, Michele Parrinello, Thomas E Markland, and David E Manolopoulos, *Efficient stochastic thermostatting of path integral molecular dynamics.*, J. Chem. Phys., **133**, 124104, 2010.
32. Thomas E Markland and David E Manolopoulos, *A refined ring polymer contraction scheme for systems with electrostatic interactions*, Chem. Phys. Lett., **464**, 256, 2008.
33. Thomas E Markland and David E Manolopoulos, *An efficient ring polymer contraction scheme for imaginary time path integral simulations.*, J. Chem. Phys., **129**, 024105, 2008.
34. T Schneider and E Stoll, *Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions*, Phys. Rev. B, **17**, 1302–1322, 1978.
35. Hans C Andersen, *Molecular dynamics simulations at constant pressure and/or temperature*, J. Chem. Phys., **72**, 2384–2393, 1980.
36. H J C Berendsen, J P M Postma, W F Van Gunsteren, A DiNola, and J R Haak, *Molecular dynamics with coupling to an external bath*, J. Chem. Phys., **81**, 3684, 1984.

37. Shuichi Nosé, *A unified formulation of the constant temperature molecular dynamics methods*, J. Chem. Phys., **81**, 511, 1984.
38. W G Hoover, *Canonical dynamics: Equilibrium phase-space distributions*, Phys. Rev. A, **31**, 1695–1697, 1985.
39. G Bussi, D Donadio, and M Parrinello, *Canonical sampling through velocity rescaling*, J. Chem. Phys., **126**, 14101, 2007.
40. Michele Ceriotti, Giovanni Bussi, and Michele Parrinello, *Langevin Equation with Colored Noise for Constant-Temperature Molecular Dynamics Simulations*, Phys. Rev. Lett., **102**, 020601, 2009.
41. Jianshu Cao and Gregory A Voth, *The formulation of quantum statistical mechanics based on the Feynman path centroid density. IV. Algorithms for centroid molecular dynamics*, J. Chem. Phys., **101**, 6168–6183, 1994.
42. I R Craig and D E Manolopoulos, *Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics*, J. Chem. Phys., **121**, 3368, 2004.
43. Rosana Collepardo-Guevara, Ian R Craig, and David E Manolopoulos, *Proton transfer in a polar solvent from ring polymer reaction rate theory.*, J. Chem. Phys., **128**, 144502, 2008.
44. Seogjoo Jang and Gregory a. Voth, *A derivation of centroid molecular dynamics and other approximate time evolution methods for path integral centroid variables*, J. Chem. Phys., **111**, 2371, 1999.
45. Bastiaan J Braams and David E Manolopoulos, *On the short-time limit of ring polymer molecular dynamics.*, J. Chem. Phys., **125**, 124105, 2006.
46. Jeremy O Richardson and Stuart C Althorpe, *Ring-polymer molecular dynamics rate-theory in the deep-tunneling regime: Connection with semiclassical instanton theory*, J. Chem. Phys., **131**, 214106, 2009.
47. Timothy J H Hele and Stuart C Althorpe, *Derivation of a true ($t\rightarrow 0^{+}$) quantum transition-state theory. I. Uniqueness and equivalence to ring-polymer molecular dynamics transition-state-theory.*, J. Chem. Phys., **138**, 084108, 2013.
48. Seogjoo Jang, Anton V. Sinitskiy, and Gregory a. Voth, *Can the ring polymer molecular dynamics method be interpreted as real time quantum dynamics?*, J. Chem. Phys., **140**, 154103, 2014.
49. Alexander Witt, Sergei D Ivanov, Motoyuki Shiga, Harald Forbert, and Dominik Marx, *On the applicability of centroid and ring polymer path integral molecular dynamics for vibrational spectroscopy.*, J. Chem. Phys., **130**, 194510, 2009.
50. Mariana Rossi, Michele Ceriotti, and David E Manolopoulos, *How to remove the spurious resonances from ring polymer molecular dynamics.*, J. Chem. Phys., **140**, 234116, 2014.
51. Mariana Rossi, Hanchao Liu, Francesco Paesani, Joel Bowman, and Michele Ceriotti, *Communication: On the consistency of approximate quantum dynamics simulation methods for vibrational spectra in the condensed phase*, J. Chem. Phys., **141**, 181101, 2014.
52. Michele Ceriotti, David E Manolopoulos, and Michele Parrinello, *Accelerating the convergence of path integral dynamics with a generalized Langevin equation.*, J. Chem. Phys., **134**, 84104, 2011.
53. Michele Ceriotti and David E Manolopoulos, *Efficient First-Principles Calculation of the Quantum Kinetic Energy and Momentum Distribution of Nuclei*, Phys. Rev. Lett., **109**, 100604, 2012.

54. C W Gardiner, *Handbook of Stochastic Methods*, Springer, Berlin, third edition, 2003.
55. P Langevin, *The theory of Brownian movement*, CR Acad. Sci, **146**, 530, 1908.
56. Robert Zwanzig, *Nonequilibrium statistical mechanics*, Oxford University Press, New York, 2001.
57. Robert Zwanzig, *Memory Effects in Irreversible Thermodynamics*, Phys. Rev., **124**, 983–992, 1961.
58. B J Berne and Dieter Forster, *Topics in Time-Dependent Statistical Mechanics*, Annu. Rev. Phys. Chem., **22**, 563–596, 1971.
59. L Kantorovich, *Generalized Langevin equation for solids. I. Rigorous derivation and main properties*, Phys. Rev. B, **78**, 94304, 2008.
60. L. Stella, C. D. Lorenz, and L. Kantorovich, *Generalized Langevin equation: An efficient approach to nonequilibrium molecular dynamics of open systems*, Phys. Rev. B, **89**, 134303, 2014.
61. A A Maradudin, T Michel, A R McGurn, and E R Méndez, *Enhanced backscattering of light from a random grating*, Ann. Phys. (N. Y)., **203**, 255–307, 1990.
62. F Marchesoni and P Grigolini, *On the extension of the Kramers theory of chemical relaxation to the case of nonwhite noise*, J. Chem. Phys., **78**, 6287, 1983.
63. Craig C. Martens, *Qualitative dynamics of generalized Langevin equations and the theory of chemical reaction rates*, J. Chem. Phys., **116**, 2516–2528, 2002.
64. J Luczka, *Non-Markovian stochastic processes: Colored noise*, Chaos, **15**, 26107, 2005.
65. Michele Ceriotti, Giovanni Bussi, and Michele Parrinello, *Colored-Noise Thermostats à la Carte*, J. Chem. Theory Comput., **6**, 1170–1180, 2010.
66. Joseph A Morrone, Thomas E Markland, Michele Ceriotti, and B J Berne, *Efficient multiple time scale molecular dynamics: Using colored noise thermostats to stabilize resonances.*, J. Chem. Phys., **134**, 14103, 2011.
67. Michele Ceriotti, Giovanni Bussi, and Michele Parrinello, *Nuclear Quantum Effects in Solids Using a Colored-Noise Thermostat*, Phys. Rev. Lett., **103**, 30603, 2009.
68. Scott Habershon and David E Manolopoulos, *Zero point energy leakage in condensed phase dynamics: An assessment of quantum simulation methods for liquid water*, J. Chem. Phys., **131**, 244518, 2009.
69. Hichem Dammak, Yann Chalopin, Marine Laroche, Marc Hayoun, and Jean-Jacques Greffet, *Quantum Thermal Bath for Molecular Dynamics Simulation*, Phys. Rev. Lett., **103**, 190601, 2009.
70. O. N. Bedoya-Martínez, Jean-Louis Barrat, and David Rodney, *Computation of the thermal conductivity using methods based on classical and quantum molecular dynamics*, Phys. Rev. B, **89**, 014303, 2014.
71. Michele Ceriotti, *A novel framework for enhanced molecular dynamics based on the generalized Langevin equation*, PhD thesis, ETH Zürich, 2010.
72. Michele Ceriotti, Giacomo Miceli, Antonino Pietropaolo, Daniele Colognesi, Angeloclaudio Nale, Michele Catti, Marco Bernasconi, and Michele Parrinello, *Nuclear quantum effects in ab initio dynamics: Theory and experiments for lithium imide*, Phys. Rev. B, **82**, 174306, 2010.
73. Ali A Hassanali, Jerome Jérôme Cuny, Michele Ceriotti, Chris J Pickard, and Michele Parrinello, *The Fuzzy Quantum Proton in the Hydrogen Chloride Hydrates*, J. Am. Chem. Soc., **134**, 8557–8569, 2012.
74. Scott Habershon, Thomas E Markland, and David E Manolopoulos, *Competing quan-*

tum effects in the dynamics of a flexible water model., J. Chem. Phys., **131**, 24501, 2009.
75. Hsiao S Mei, Mark E Tuckerman, Diane E Sagnella, and Michael L Klein, *Quantum Nuclear ab Initio Molecular Dynamics Study of Water Wires*, J. Phys. Chem. B, **102**, 10446–10458, 1998.
76. P L Geissler, C Dellago, D Chandler, J Hutter, and M Parrinello, *Autoionization in liquid water*, Science, **291**, 2121–2124, 2001.
77. Mark E Tuckerman, Dominik Marx, and Michele Parrinello, *The nature and transport mechanism of hydrated hydroxide ions in aqueous solution.*, Nature, **417**, 925–9, 2002.
78. G. Reiter, C. Burnham, D. Homouz, P. Platzman, J. Mayers, T. Abdul-Redah, A. Moravsky, J. Li, C.-K. Loong, and A. Kolesnikov, *Anomalous Behavior of Proton Zero Point Motion in Water Confined in Carbon Nanotubes*, Phys. Rev. Lett., **97**, 247801, 2006.
79. Giovanni Romanelli, Michele Ceriotti, David E Manolopoulos, Claudia Pantalei, Roberto Senesi, and Carla Andreani, *Direct Measurement of Competing Quantum Effects on the Kinetic Energy of Heavy Water upon Melting*, J. Phys. Chem. Lett., **4**, 3251–3256, 2013.
80. R. A. Berner, S. T. Petsch, J. A. Lake, D. J. Beerling, B. N. Popp, R. S. Lane, E. A. Laws, M. B. Westley, N. Cassar, F. I. Woodward, and W. P. Quick, *Isotope Fractionation and Atmospheric Oxygen: Implications for Phanerozoic O2 Evolution*, Science, **287**, 1630–1633, 2000.
81. T E Markland and B J Berne, *Unraveling quantum mechanical effects in water using isotopic fractionation*, Proc. Natl. Acad. Sci. USA, **109**, 7988–7991, 2012.
82. Yuki Nagata, Ruben E Pool, Ellen H G Backus, and Mischa Bonn, *Nuclear Quantum Effects Affect Bond Orientation of Water at the Water-Vapor Interface*, Phys. Rev. Lett., **109**, 226101, 2012.
83. Jian Liu, Richard S. Andino, Christina M. Miller, Xin Chen, David M. Wilkins, Michele Ceriotti, and David E. Manolopoulos, *A Surface-Specific Isotope Effect in Mixtures of Light and Heavy Water*, J. Phys. Chem. C, **117**, 2944–2951, 2013.
84. Lu Wang, Michele Ceriotti, and Thomas E. Markland, *Quantum fluctuations and isotope effects in ab initio descriptions of water*, J. Chem. Phys., **141**, 104502, 2014.
85. X.-Z. Li, B Walker, and A Michaelides, *Quantum nature of the hydrogen bond*, Proc. Natl. Acad. Sci. USA, **108**, 6369–6373, 2011.
86. Ross H McKenzie, Christiaan Bekker, Bijyalaxmi Athokpam, and Sai G Ramesh, *Effect of quantum nuclear motion on hydrogen bonding.*, J. Chem. Phys., **140**, 174508, 2014.
87. Michele Ceriotti, Jérôme Cuny, Michele Parrinello, and David E Manolopoulos, *Nuclear quantum effects and hydrogen bond fluctuations in water.*, Proc. Natl. Acad. Sci. USA, **110**, 15591–6, 2013.
88. Michele Ceriotti, Joshua More, and David E. Manolopoulos, *i-PI: A Python interface for ab initio path integral molecular dynamics simulations*, Comput. Phys. Commun., **185**, 1019–1026, 2014.
89. Ali Hassanali, Meher K Prakash, Hagai Eshet, and Michele Parrinello, *On the recombination of hydronium and hydroxide ions in water.*, Proc. Natl. Acad. Sci. USA, **108**, 20410–5, 2011.

90. Glenn J Martyna, Adam Hughes, and Mark E Tuckerman, *Molecular dynamics algorithms for path integrals at constant pressure*, J. Chem. Phys., **110**, 3275, 1999.
91. "CP2K", http://www.cp2k.org.
92. Joost VandeVondele, Matthias Krack, Fawzi Mohamed, Michele Parrinello, Thomas Chassaing, and Jürg Hutter, *Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach*, Comput. Phys. Commun., **167**, 103–128, 2005.
93. Steve Plimpton, *Fast Parallel Algorithms for Short-Range Molecular Dynamics*, J. Comput. Phys., **117**, 1–19, 1995.
94. Volker Blum, Ralf Gehrke, Felix Hanke, Paula Havu, Ville Havu, Xinguo Ren, Karsten Reuter, and Matthias Scheffler, *Ab initio molecular simulations with numeric atom-centered orbitals*, Comput. Phys. Commun., **180**, 2175–2196, 2009.

Ab initio Molecular Dynamics Approach to Solvation and Chemical Reactions

Rodolphe Vuilleumier

Ecole Normale Supérieure-PSL Research University, Département de Chimie
24, rue Lhomond, 75005 Paris, France
and
Sorbonne Universités, UPMC Univ Paris 06, PASTEUR, 4 place Jussieu, F-75005, Paris, France
and
CNRS, UMR 8640 PASTEUR, F-75005, Paris, France

E-mail: rodolphe.vuilleumier@ens.fr

Ab initio molecular dynamics aims at simulating simultaneously nuclei and electrons. Because of the very different time-scales for the electrons and the nuclei dynamics, ab initio molecular dynamics, in its most common form, makes an adiabatic separation of the nuclei and electron systems, with the electrons being held in the electronic ground state along the dynamics. Forces on the nuclei are then computed from this electronic ground state calculation at each time-step of the simulation. This allows for the description of materials in a wide range of situations. Ab initio molecular dynamics has then been widely used for the study of solvation processes and chemical reactions in solution. At the same time, knowledge of the electronic structure along the dynamics allows direct comparison with spectroscopy experiments and to investigate the nature of the interactions. In these lectures notes, we will briefly review density functional theory based molecular dynamics and discuss how it can be applied to the study of chemical processes in solutions.

1 Introduction

Cohesion of matter relies essentially on the properties of the electronic cloud in which nuclei evolve. Computer simulations of materials thus aim at describing nuclei and electrons altogether. Ab initio molecular dynamics precisely aims at treating condensed phase systems by describing simultaneously the nuclei dynamics and the electronic structure[1–5]. While it is widely recognised that the simulation of chemical reactions, because chemical bonds are broken and newly formed, calls for an adequate description of the electronic structure, it is less often recognised that this is also the case if one also wants to simulate solvation processes such as the solvation of metallic ions or even organic molecules. Metallic ions form complexes whose structure is determined by their electronic structure. Ag^+ and Na^+ have trivially the same charge, they also have the same ionic radius. A description of these two ions taking into account only these two quantities would then lead to similar, or even identical, solvation shells while their coordination number is notably different in liquid water[6]. Such specific effect is also at the origin of poisoning by heavy metals[7]. Ab initio molecular dynamics, because it combines the explicit description of the electronic structure and of the thermal motion of the atoms, provides a natural route for the simulation of these phenomena. Moreover, the electronic structure computed along the molecular dynamics trajectory can lead on one hand to some new insight in the chemical processes occurring and on the other hand to the evaluation of electronic observables

allowing for a unique connection with spectroscopy experiments. Combination of simulation and spectroscopy experiments allow for the characterisation of the structure and species involved, which leads to the identification and understanding of the most important interactions. It is then possible to build relevant simplified models. To this respect, recent developments have allowed for the evaluation of thermodynamical quantities so that we have gone from qualitative pictures to quantitative models. The purpose of these lecture notes is to introduce the reader to ab initio molecular dynamics for the simulation of solvation and chemical reactions, and the calculation of spectroscopic and thermodynamical quantities. They are organised as follows: First, in the next section, a brief account of the ab initio molecular dyanmics technique will be given. In the second section, we will give examples of applications to solvation and to the interpretation of spectroscopy experiments of different nature as well as to chemical reactions. Finally, some extensions and routes for future developments will be discussed followed by some conclusions.

2 Ab initio Molecular Dynamics

Ab initio molecular dynamics (AIMD) describes simultaneously the time evolution of the electronic structure and of the nuclei geometry.[2–5,8] Electrons move on a much faster time scale than nuclei and AIMD is one of the first example of multiscale dynamics. In the vast majority of cases, AIMD is based on a complete, adiabatic, separation between the fast degrees of freedom, the electrons, and the slow degrees of freedom, the nuclei. The dynamics of the slowest degrees of freedom, the nuclei, is driven by an equilibrium distribution of the fast degree of freedom dynamics at fixed nuclei position. More precisely, in the Born-Oppenheimer approximation[9], the electronic state of the system is pinned onto the electronic ground state at fixed nuclear position. The nuclei then evolve on a potential energy surface that is the electronic ground state energy at each nuclear position.

In this section, we review the Born-Oppenheimer approximation that separates the electronic and nuclear degrees of freedom and then discuss the electronic structure problem for the description of liquid state.

2.1 The Born-Oppenheimer Approximation

The common presentation of the Born-Oppenheimer approximation in fact follows the derivation from Born and Huang[10], which we here summarise. We consider a composite system consisting of ions and N electrons. The instantaneous quantum state of such system is described by the system wavefunction $\Psi_S(\mathbf{R}_I, \mathbf{r}_1, \ldots, \mathbf{r}_N)$ which is an integrable function of the nuclei and electrons coordinates, $\mathbf{R}_I$ and $\mathbf{r}_1, \ldots, \mathbf{r}_N$ respectively. The dynamics of the wavefunction Ψ_S is governed by the Hamiltonian of the system which we write as:

$$\hat{H}_T = \hat{T}_N + \hat{V}_{NN} + \hat{H}, \tag{1}$$

where $\hat{T}_N$ is the kinetic energy operator for the nuclei, $\hat{V}_{NN}$ is the potential energy operator acting on nuclei degrees of freedom only and arising from the nuclei-nuclei interaction and $\hat{H}$,

$$\hat{H} = \hat{T} + \hat{V}_{ee} + \hat{V}_{Ne}, \tag{2}$$

is the electrons Hamiltonian, including the interaction between the electrons and the nuclei.

Born and Huang proposed to write the wavefunction of the full system as

$$\Psi_S(\mathbf{R}_I, \mathbf{r}_1, \ldots, \mathbf{r}_N) = \sum_i \chi_i(\mathbf{R}_I, t)\varphi_i(\mathbf{r}_1, \ldots, \mathbf{r}_N; \mathbf{R}_I), \tag{3}$$

where $\varphi_i(\mathbf{r}_1, \ldots, \mathbf{r}_N; \mathbf{R}_I)$ is the i^{th} eigenstate of the electronic hamiltonian at fixed nuclei positions $\mathbf{R}_I$. These states are normalized

$$\langle\varphi_i|\varphi_i\rangle_{\mathbf{R}_I} = \int \cdots \int |\varphi_i(\mathbf{r}_1, \ldots, \mathbf{r}_N; \mathbf{R}_I)|^2 \, \mathrm{d}^3\mathbf{r}_1 \ldots \mathrm{d}^3\mathbf{r}_N = 1 \tag{4}$$

for every nuclear position $\mathbf{R}_I$. $\chi_i(\mathbf{R}_I)$ then represents the part of the nuclear wavefunction evolving of that electronic state. The norm of these nuclear functions indicate the probability to find the electronic state in a given eigenstate.

By applying the time-dependent Schrödinger equation on this wavefunction ansatz and projecting on the electronic state i, Born and Huang obtain the coupled evolution equation of the nuclear wavefunctions:

$$\begin{aligned} i\hbar\frac{\partial|\chi_i(t)\rangle}{\partial t} = &\left(\hat{T}_N + \hat{V}_{NN}\right)|\chi_i(t)\rangle \\ &+ \langle\varphi_i|\hat{H}|\varphi_i\rangle_{\mathbf{R}_I} \times |\chi_i(t)\rangle \\ &+ \frac{1}{2M_I}\sum_j \langle\varphi_i|\nabla^2_{\mathbf{R}_I}|\varphi_j\times\rangle_{\mathbf{R}_I}|\chi_j(t)\rangle \\ &+ \frac{1}{M_I}\sum_j \langle\varphi_i|\nabla_{\mathbf{R}_I}|\varphi_j\rangle_{\mathbf{R}_I} \times \nabla_{\mathbf{R}_I}|\chi_j(t)\rangle. \end{aligned} \tag{5}$$

The last two terms couple the nuclear dynamics on different electronic states. In the limit $m_e/M_I \to 0$ these terms are expected to be small. Neglecting these terms is the essence of the Born-Oppenheimer approximation. With this approximation, the nuclear dynamics on each electronic state is decoupled from that on other electronic state and can be described independently. At thermal equilibrium, only the electronic ground state is significantly populated in the vast majority of situations and the nuclear evolution on the electronic ground state is then described by the nuclear time-dependent Schrödinger equation:

$$\begin{aligned} i\hbar\frac{\partial|\chi_0(t)\rangle}{\partial t} &= \left(\hat{T}_N + \hat{V}_{NN}\right)|\chi_0(t)\rangle + \epsilon_0(\mathbf{R}_I)|\chi_0(t)\rangle \\ &= \left(\hat{T}_N + \hat{V}_{BO}\right)|\chi_0(t)\rangle, \end{aligned} \tag{6}$$

where $\epsilon_0(\mathbf{R}_I) = \langle\varphi_0|\hat{H}|\varphi_0\rangle_{\mathbf{R}_I}$ is the electronic groundstate energy. The nuclei thus evolve on a potential energy surface $V_{BO}(\mathbf{R}_I) = V_{NN}(\mathbf{R}_I) + \epsilon_0(\mathbf{R}_I)$ sum of the nuclei-nuclei interaction (Coulomb repulsion) and the electronic ground state energy at fixed nuclear position $\mathbf{R}_I$.

The evolution equation for the nuclei has the usual structure of a Schrödinger equation of a set of particles subject to a potential energy surface. In molecular dynamics simulation, we take the classical limit for the nuclei dynamics, which then evolve according to the forces $\mathbf{F}_I$:

$$\mathbf{F}_I = -\frac{\partial V_{BO}(\mathbf{R}_I)}{\partial \mathbf{R}_I} = -\frac{\partial V_{NN}(\mathbf{R}_I)}{\partial \mathbf{R}_I} - \frac{\partial \epsilon_0(\mathbf{R}_I)}{\partial \mathbf{R}_I}. \tag{7}$$

Because of the normalization of the electronic wavefunction φ_0, the calculation of the nuclear force is simplified using the Hellman-Feynman theorem, which expresses that

$$\frac{\partial \epsilon_0(\mathbf{R}_I)}{\partial \mathbf{R}_I} = \langle \varphi_0 | \frac{\partial \hat{H}}{\partial \mathbf{R}_I} | \varphi_0 \rangle_{\mathbf{R}_I}. \tag{8}$$

It is then not necessary to know the response of the electronic wavefunction to a nuclear displacement to evaluate the nuclear forces. The classical force felt by the nuclei are then the quantum average of the force operator on the electronic ground state.

This constitutes the basis of ab initio molecular dynamics: we simulate the evolution of a set of nuclei whose forces are obtained from the evaluation of the electronic ground state energy and its gradient with respect to the nuclear positions. Classical molecular dynamics simulations then use a model of these forces instead of an explicit evaluation of the electronic ground state energy. However, the electronic state calculation is computationally very expensive and to achieve large system sizes for describing condensed phase systems such as solvation, Density Functional Theory (DFT) is employed in most cases, and we give below a brief account of the DFT theory. Note however, that semi-empirical, tight-binding, methods may be used and recently liquid water was simulated using a description of the electronic state at the second-order Møller-Plesset level (MP2). However, at the moment, only Monte-Carlo simulations at the MP2 level are accessible.

2.2 Density Functional Theory

The difficulty in solving for the ground state electronic wavefunction at a given ionic configuration arises from the fact that electrons are interacting fermions. The Hamiltonian of an N electrons system in the external field $v_{ext}(\mathbf{r})$ created by the nuclei is:

$$\hat{H} = \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} \tag{9}$$

where $\hat{T}$ is the kinetic energy operator:

$$\hat{T} = -\frac{1}{2} \sum_{i=1}^{N} \hat{\nabla}_i^2, \tag{10}$$

and $\hat{V}_{ee}$ represents the electron-electron repulsion:

$$\hat{V}_{ee} = \frac{1}{2} \sum_{i \neq j} \frac{1}{|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j|} \tag{11}$$

($\hat{\mathbf{r}}_i$ is the position operator for electron i).

The electronic ground state $\Psi_0(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ is a function of $3N$ variables which satisfies the fermion antisymmetry property:

$$\Psi_0(\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_N) = -\Psi_0(\mathbf{r}_1, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_N). \tag{12}$$

Ψ_0 is the eigenvector of $\hat{H}$ associated with the lowest eigenvalue E_0:

$$\hat{H}\Psi_0(\mathbf{r}_1, \ldots, \mathbf{r}_N) = E_0 \Psi_0(\mathbf{r}_1, \ldots, \mathbf{r}_N) \tag{13}$$

or, alternatively, Ψ_0 can be obtained from a variational principle:

$$E_0 = \min_{\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)} \langle \Psi | \hat{H} | \Psi \rangle, \tag{14}$$

the minimization of $\langle\Psi|\hat{H}|\Psi\rangle$ being realized for $\Psi = \Psi_0$ (non-degenerate ground state).

Since the electrons are interacting the ground state wavefunction Ψ_0 in general can not be factorized in a single Slater determinant:

$$\psi_{Slater} \equiv \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \dots & \phi_N(\mathbf{r}_1) \\ \vdots & & & \vdots \\ \phi_1(\mathbf{r}_N) & \dots & \dots & \phi_N(\mathbf{r}_N) \end{vmatrix}, \tag{15}$$

where ϕ_i are one-electron orbitals, which does not exhibit any electron-electron correlation. Thus the electronic wavefunction is a very complicated object that can not even be stored for more than a few electrons. It is for example the purpose of Quantum Monte Carlo techniques to try to perform calculations (calculating the expectation value of the Hamiltonian) using directly a representation of the electronic wavefunction[11,12].

The strength of Density Functional Theory (DFT)[13–15] is based on the fact that the electronic wavefunction $\Psi_0(\mathbf{r}_1, \dots, \mathbf{r}_N)$ of the electronic *ground state* of the system can be entirely described only by its electron density $n^0(\mathbf{r})$, as stated by the Hohenberg-Kohn theorem[16]. It is based on a *minimization* principle stating that the ground state electronic density minimizes an *energy functional*. Thus in principle we have to consider and manipulate a much simpler object, the electronic density, which is simply a function of $\mathbb{R}^3$.

The Hohenberg-Kohn theorem[16,13–15] states that there exists a one to one map between external potentials and the ground state electronic density:

$$n^0(\mathbf{r}) \to v(\mathbf{r}). \tag{16}$$

For an external potential $v(\mathbf{r})$ there is a unique ground-state wavefunction Ψ_0[a] and as a result it gives rise to a unique ground state density $n^0(\mathbf{r})$. The Hohenberg-Kohn theorem states that for electronic densities $n(\mathbf{r})$ which are v-representable, that is which are the ground state of some external potential, this external potential which gives rise to them is unique. This can be expressed by the following map:

$$\underbrace{v(\mathbf{r}) \quad \to \quad \Psi_0 \quad \to \quad n^0(\mathbf{r})}_{\text{HK theorem}} \tag{17}$$

Knowledge of $n(\mathbf{r})$ gives knowledge of $v(\mathbf{r})$ and of Ψ_0, thus all properties of the electronic system and in particular any expectation value of an observable $\hat{O} = \langle\Psi|\hat{O}|\Psi\rangle$, are functionals of $n(\mathbf{r})$ ($O[n]$ for the observable $\hat{O}$).

The total energy of the system is then itself a *functional* $E[n]$ of the ground state electronic density[16,13]. Separating the interaction with the external potential from the rest (kinetic energy and electron-electron interaction) we can write the total energy as

$$E[n] = \int d^3\mathbf{r}\, n(\mathbf{r})v(\mathbf{r}) + F_{HK}[n] \tag{18}$$

[a] For *non-degenerate* ground states, which is the very general case, degenerate systems are to be treated particularly.

which defines the Hohenberg-Kohn functional $F_{HK}[n]$ as

$$F_{HK}[n] = T[n] + V_{ee}[n] \tag{19}$$

$$T[n] = \langle \Psi[n] | \hat{T} | \Psi[n] \rangle \tag{20}$$

$$V_{ee}[n] = \langle \Psi[n] | \hat{V}_{ee} | \Psi[n] \rangle, \tag{21}$$

where $\psi[n]$ is the ground state wavefunction, functional of n. Note that in eq. (18) v is to be understood as a functional $v[n]$ of n through the map (17). We will note that[17] $F_{HK}[n] = \min_{\Psi \to n} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle$, realised for $\Psi = \Psi[n]$, the ground state associated with the electron density n.

In most cases, what is given to us is the external potential $v_0(\mathbf{r})$ instead of the ground state density. In its turn, the ground state density n_0 for the external potential v_0 can be shown to minimize the energy functional[17]:

$$E_{v_0}[n] = \int d^3\mathbf{r}\, n(\mathbf{r}) v_0(\mathbf{r}) + F_{HK}[n] \tag{22}$$

in which v_0 is fixed, while $F_{HK}[n]$ is the universal functional defined above (independent on v_0).

These theorems are proof ox existence of an energy functional but do not give a way to construct it and approximations have to be made. The most difficult term to express as a functional of $n(\mathbf{r})$ is the kinetic energy term $T[n]$. Kohn and Sham proposed to reintroduce electronic orbitals to evaluate this term, but to keep the problem tractable these are orbitals of an auxiliary non-interacting electron system[18, 13–15].

The ground state wavefunction Ψ_s of a system of such an N non-interacting electron system in an external potential $v_s(\mathbf{r})$ is a single slater determinant

$$\Psi_s = \frac{1}{\sqrt{N!}} \det[\phi_1 \phi_2 \dots \phi_N] \tag{23}$$

made of the N orbitals ϕ_i which are the N lowest eigenvectors of the single electron Hamiltonian $\hat{H}_S$:

$$\hat{H}_s = -\frac{1}{2}\nabla^2 + v_s(\mathbf{r}). \tag{24}$$

That is the orbitals ϕ_i satisfy

$$\hat{H}_s \phi_i = \epsilon_i \phi_i \text{ with } \langle \phi_i | \phi_j \rangle = \delta_{ij}. \tag{25}$$

The ground state density of this non-interacting electron system is

$$n_s(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2. \tag{26}$$

If for a given electronic density $n(\mathbf{r})$ there exists an external potential v_s whose non-interacting electron ground state density is equal to n:

$$n_s(\mathbf{r}) \equiv n(\mathbf{r}) \tag{27}$$

then n is said to be non-interacting v-representable.

The Hohenberg-Kohn theorem is valid also for non-interacting systems such that there exists a one to one map between n and v_s; the map

$$n(\mathbf{r}\,) \rightleftarrows v_s(\mathbf{r}\,) \rightleftarrows \{\phi_i(\mathbf{r}\,)\} \tag{28}$$

is unique. The Hohenberg-Kohn functional for the non-interacting electron system is simply the kinetic energy T_s:

$$T_s[n] = \sum_{i=1}^{N} \langle \phi_i[n] | -\frac{1}{2}\nabla^2 | \phi_i[n] \rangle. \tag{29}$$

We will then approximate the kinetic energy functional $T[n]$ for the full system by the kinetic energy functional $T_s[n]$ of this auxiliary non-interacting electron system.

The total energy of the system can be decomposed into[18]

$$E[n] = T_s[n] + \int d^3\mathbf{r}\, n(\mathbf{r}\,) v(\mathbf{r}\,) + J[n] + E_{xc}[n], \tag{30}$$

where $J[n]$ is the Hartree term:

$$J[n] = \frac{1}{2} \iint d^3\mathbf{r} d^3\mathbf{r}\,' \frac{n(\mathbf{r}\,) n(\mathbf{r}\,')}{|\mathbf{r} - \mathbf{r}\,'|} \tag{31}$$

and equation (30) defines the exchange-correlation energy $E_{xc}[n]$ as the remainder:

$$E_{xc}[n] = (T[n] - T_s[n]) + (V_{ee}[n] - J[n])\,. \tag{32}$$

The Hartree term $J[n]$ describes the electrostatic energy of the electronic system within a mean field approximation: it is the classical electrostatic energy of a charge density $n(\mathbf{r}\,)$. The exchange correlation energy is mainly the difference between the true electrostatic energy $V_{ee}[n]$ and the mean field term $J[n]$ but also contains a contribution from the kinetic energy, $T[n] - T_s[n]$. Once again, the Hohenberg-Kohn theorem for interacting and non-interacting electrons proves the existence and the uniqueness of the functional $E_{xc}[n]$ but this functional is still unknown to us. It has proved however much easier to find approximations to $E_{xc}[n]$ than to the Hohenberg-Kohn functional $F_{HK}[n]$ (18). There exists a whole Jacob's ladder of approximations for the exchange and correlation energy, from local density approximation (LDA), generalized gradient approximation (GGA), hybrid, double hybrid, range separated functionals etc. It is not the purpose of these lecture notes to give an account of these methods and the reader is invited to refer to the literature for that purpose[19–22]. We will only briefly mention some points relative to the AIMD description of liquids, but we will first describe some implementations of AIMD.

2.3 Basis sets

To perform actual computation of the Kohn-Sham orbitals, one need to define a representation, that is a choice of basis-set that is used to define the Kohn-Sham orbitals as linear combinations of the basis-set elements. The Kohn-Sham orbitals $\{\phi_i\}$ are expanded on a basis set $\{\varphi_\alpha\}$:

$$\phi_i(\mathbf{r}\,) = \sum_\alpha c_i^\alpha \varphi_\alpha(\mathbf{r}\,).$$

In this section we will quickly discuss the special case of plane wave basis set, as they have been often used in first-principle Molecular Dynamics simulations.

The typical basis sets are plane waves and atomic centered basis sets. For plane waves, the basis set elements are defined as[2,23]

$$\varphi_\alpha(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} e^{-i\mathbf{G}_\alpha \cdot \mathbf{r}} \tag{33}$$

where Ω is the volume of the simulation box. For an orthorhombic box with lengths L_x, L_y and L_z, the wavevectors $\mathbf{G}$ are

$$\mathbf{G} = i \cdot \frac{2\pi}{L_x} \cdot \vec{x} + j \cdot \frac{2\pi}{L_y} \cdot \vec{y} + k \cdot \frac{2\pi}{L_z} \cdot \vec{z}\,; \text{with } i,j,k \in \mathbb{Z}. \tag{34}$$

The main advantage of a plane wave basis set, in view of Molecular Dynamics, is the independence of the basis set elements with respect to the ionic positions.[2] As a result, the Hellmann-Feynman theorem can be applied straightforwardly, without additional so-called Pulay terms arising from a basis set that would be dependent on the nuclei positions. The forces on the ions will be calculated at virtually no extra-cost. There is also no Basis Set Superposition Error for the same reasons. Another advantage of plane wave basis sets is that their quality depends only on the number of wave-vectors considered. The number of basis set elements is controlled by an energy cutoff E_{cut} such that all plane wave whose kinetic energy is lower than this value are included:

$$E_{kin} = \frac{1}{2}\|\mathbf{G}\|^2 < E_{cut}. \tag{35}$$

It is thus easier both to compare results and to make convergence studies with only one number defining the quality of the basis set. Finally, on the computional side, plane wave basis sets have the further advantage of being orthonormal. As a result also, they can not become over-complete (one element being exactly or even approximately a linear combination of other elements, leading to divergences).

However, plane wave basis sets also have disadvantages. The first one is probably the very large number of basis set elements which can range from a few 10000 to a few 10^6. To avoid this number of basis set elements to become even higher so that calculations would become untractable it is absolutely necessary to employ pseudopotentials: Only valence electrons are considered, not core electrons; it results from this that the electron-ion interactions are not simply the fundamental coulomb attraction.[24–33] This interaction also loses its local character (except for very simple cases such as the hydrogen atom) and takes the general form

$$\hat{V}^{ps} = V_{local}(\hat{r}) + \sum_{kl} \hat{P}_k h_{kl} \hat{P}_l, \tag{36}$$

where $V_{local}(\hat{r})$ is an isotropic local potential and $\hat{P}_l$ are projectors constructed from atomic orbitals (h_{kl} are parameters, real numbers).

Lately, atom centered basis set have become more popular in the framework of ab initio molecular dynamics. Atom centered basis-sets in general contain much less elements that plane waves. Pseudopotentials may still be used, and are most often used, in order to lower the cost of AIMD. Also, atome centered basis sets are localized in space, which allow for the development of order-N methods, at least for the construction of the Kohn-Sham

hamiltonian. Order-N methods have the desired property of having a numerical cost that scales linearly with the system size. CP2K[34] uses gaussian basis sets, which have the advantage that the kinetic energy, an overlap can be computed analytically. SIESTA for example uses trully localized basis-sets that are strictly zero outside a finite radius[35].

Because of the central role of the electronic density in DFT, one also need a representation for it. With plane wave basis-sets, it is natural to similarly express the electronic density as a sum of plane waves. In part because of the common use of periodic boundary conditions in AIMD, plane wave representations of the electronic density are also employed with atom centered basis sets.

2.4 Car-Parrinello Molecular Dynamics and Born-Oppenheimer Molecular dynamics

Having a way to describe the electronic state, one can then compute it along the nuclei dynamics as needed for AIMD. In principle, the Kohn-Sham orbitals are obtained by diagonalization of the non-interacting electrons hamiltonian in an external potential $v_s(\mathbf{r})$, the Kohn-Sham hamiltonian, which itself depends on the Kohn-Sham orbitals and a self-consistency condition defines the solution of the Kohn-Sham problem. In 1985, Car and Parrinello[1] expressed this condition as a variational problem on the set of Kohn-Sham orbitals and showed that the solution can then be obtained by minimization techniques, such as steepest descent or conjugate gradients. Using a plane wave basis set to describe the Kohn-Sham orbitals in periodic boundary conditions they could efficiently compute the electronic state of eight silicon atoms in LDA. The next big step they made was to state that the Kohn-Sham orbitals can be propagated along with the nuclei as long as this propagation ensures an adiabatic separation of time scales between the Kohn-Sham orbitals and the nuclear motion. However, this separation of time scales does not have to be as large as for the actual electronic system. By making the time scale gap smaller, the numerical propagation would be made easier. To achieve this they proposed an extended Lagrangian, now known as the Car-Parrinello Lagrangian, to describe the system dynamics, including the Kohn-Sham orbitals as explicit degrees of freedom. This Lagrangian can be written as

$$\mathcal{L}(\mathbf{R},\dot{\mathbf{R}},\phi,\dot{\phi}) = \frac{1}{2}\mu\sum_i\langle\dot{\phi}_i|\dot{\phi}_i\rangle + \frac{1}{2}\sum_I M_I\dot{\mathbf{R}}_I^2 - E[\mathbf{R},\phi] + \sum_{ij}\Lambda_{ij}(\langle\phi_i|\phi_i\rangle - \delta_{ij}). \quad (37)$$

In this expression $E[\mathbf{R},\phi]$ is the Kohn-Sham energy for a given set of nuclear coordinates and Kohn-Sham orbitals, and we see that the Kohn-Sham energy is expressed as a functional of the Kohn-Sham orbitals instead of a functional of the electronic density: this is possible since the two have the same minimum. The last term in the Car-Parrinello Lagrangian are Lagrange multipliers to ensure the orthogonality of the Kohn-Sham orbitals during the dynamics. μ is a fictitious mass for the electronic system: for small enough μ we have an adiabatic separation of the Kohn-Sham orbitals dynamics and the nuclear dynamics. The nuclear dynamics is then driven by an average force over a distribution of the Kohn-Sham orbitals close to the minimum of the Kohn-Sham energy at fixed nuclear positions: this is precisely the Hellman-Feynman force that drives the nuclear dynamics in the Born-Oppenheimer approximation. Small fictitious masses μ however make the typical time-scale for the Kohn-Sham orbitals smaller and small timesteps have to be used, making

the computation more costly. A compromised has thus to be found between separation of time-scales and computational cost[5].

This was really the start of ab initio Molecular Dynamics as implemented in popular codes such as CPMD[36], Quantum Espresso[37] and others. On the methodological side, developments have been carried out on more efficient minimisation techniques[38] and it is now possible to carry out genuine Born-Oppenheimer dynamics, where the Kohn-Sham problem is first solved at fixed nuclear positions, then the Hellman-Feynamn forces are computed to propagate the nuclei dynamics[34]. Efficiency in part relies in on the development of efficient methods for extrapolating the Kohn-Sham orbitals along the dynamics to serve as good initial guesses for the DFT calculation[39,40]. The Born-Oppenheimer propagation then fully parallels a force-field molecular dynamics where the forces are calculated in a peculiar way. The timestep are then very similar, of the order of 1 fs, while the fast, although fictitious, time-scale of the Car-Parrinello propagation calls for much shorter timesteps, of the order of 0.05 fs. Also, the choice of basis set may influence the choice of the propagation method. Plane wave, including a large number of basis set elements, which are however orthogonal and independent of the nuclei positions, are very efficient for the Car-Parrinello propagation. In particular, the calculation of the forces on the atoms comes at nearly no extra cost, such that the necessary use of a short timestep is not a problem. With atomic center basis sets, force calculation are more expensive, while extrapolation methods are more efficient and Born-Oppenheimer propagation is favoured in this case.

3 Applications to Solvation and Chemical Reactions

3.1 Liquid water

The proper description of liquid water by AIMD and more generaly of molecular systems, thanks to the development of GGA functionals, made AIMD applicable to the study of chemistry. However, the agreement with experiment is not perfect. To some extend it is lower than one can be achieved by optimizing an empirical force-field, but the versatility of AIMD has no equivalent. AIMD water is known to be too structured at room temperature and experimental water density[41,42]. It results in a too low self-diffusion constant. The phase diagram of AIMD water is also displaced with a too low temperature of the critical point[43] and a too high melting temperature[44]. The relaxed density at ambient temperature and ambient pressure is also nearly 20% too low[45]. The main origin of these discrepancies with experiment has been identified as being the lack of long-range dispersion, London forces scaling as $1/R^6$ for an interatomic distance R, in GGA functionals, which rely only on a local description of the electronic density. It is nowadays possible to account for the long range dispersion[46–48] and inclusion of these long range effects make the agreement much better with experimental data[45,49–51]. New developments in the use of more advanced electronic state methods, such as hybrid functional or MP2 calculation, make this agreement quite remarkable[52]. Nevertheless, GGA's corrected for dispersion give very good results for a large class of closed-shell, molecular or ionic, liquids. Also, to many respect, AIMD water contains most of the important chemistry of water, with a proper description of the modulation of the hydrogen bonds by the local structure, the discrepancies coming rather from the non-hydrogen bonded configurations.

3.2 Structure of solvated species – spectroscopy

This successful description of water paved the way to the study of solvated species. Many ions have now been studied through AIMD and also molecules[6,53–77]. As mentioned in the introduction, the solvation of metallic cations attracted a lot of attention, from transition metals[78–81], to coinage metals[6,69,82–87], lanthanides[88,89], actinides[90] and other heavy metals. Simulations were performed not only in ambient water but also at extreme conditions[91,84–87]. AIMD makes it indeed possible to study conditions of high pressure and high temperature because it is not bound to an empirical potential fitted to a specific condition. Because the dielectric constant of water goes down tremendously as a function of temperature, ions form more stable complexes and AIMD was able to study complexes of gold and silver to help understand the mobilisation of these species in hydrothermal sources[84–87]. At ambient conditions, stable complexes are formed with multivalent cations[92,93,89]. There too AIMD made it possible to study the structure, stability and also acidity of such complexes.

Spectroscopy then plays an important role in this field as it enables a direct comparison of the solvation structures with experiment. EXAFS is of course often used as it is directly sensitive to the local structure of the ions, but it is not very sensitive to the three dimensional solvation structure and to the total coordination number[90,85]. Near-edge x-ray absorption spectroscopies[94,95] can then help to this respect. Also, nuclear magnetic resonance[96–98] and electron paramagnetic resonance[99], UV-vis absorption[100,83,71,101], and nowadays valence photoelectron spectroscopy[102,103], for which experimental data are now available, are very valuable because they probe directly the interaction of the solute with water, through the induced modification of the electronic structure.

Among these spectroscopies, vibrational spectrosc opy has a special place. It probes the solvent but also the structure of molecular solute and their interaction with water. In molecular dynamics simulations, the infrared spectrum is readily available through the fluctuation-dissipation theorem that expresses the signal as[104]

$$n(\omega)\alpha(\omega) = \frac{\beta\pi}{3cV\epsilon_0}\int_{-\infty}^{+\infty} e^{i\omega t}\langle\vec{M}(0)\vec{M}(t)\rangle_T \,\mathrm{d}t, \tag{38}$$

where $n(\omega)$ and $\alpha(\omega)$ are the refractive index and linear absorption coefficient (Beer-Lambert coefficient, respectively. $\vec{M}(t)$ is system polarization at time t and $\langle\cdot\rangle_T$ designates an average at temperature T. $\beta = \frac{1}{k_B T}$ with k_B the Boltzmann constant, c is the speed of light and V the system volume. The infrared signal can then be calculated from the data of the polarisation $\vec{M}(t)$ of the simulated system along the trajectory. However, this pauses a problem with periodic boundary conditions since the position operator on the electronic system is hill defined. Resta[105,106] proposed, using the Berry phase approach[107–109], to define the expectation value of the position operator in PBC as (here for the x axis only):

$$\langle x\rangle = \frac{L}{2\pi}\Im\ln\langle\Psi|e^{i\frac{2\pi}{L}\hat{x}}|\Psi\rangle. \tag{39}$$

Note that $\langle x\rangle$ is then defined modulo L.

$\langle\Psi|e^{i\frac{2\pi}{L}\hat{x}}|\Psi\rangle$ is a many-body quantity as $|\Psi\rangle$ is the all-electron wavefunction and the operator $e^{i\frac{2\pi}{L}x}$ is not the sum of one-electron operators, but the product. In the Kohn-Sham method we will replace $|\Psi\rangle$ by the wavefunction of the auxiliary non-interacting electron

system $|\Psi_s\rangle$ which is a single Slater determinant constructed from the minimal Kohn-Sham orbitals:

$$\Psi_s = \frac{1}{\sqrt{N!}} \det[\phi_1 \phi_2 \dots \phi_N]. \tag{40}$$

In that case the expectation value of the position operator takes the form:

$$\langle x \rangle = \frac{L}{2\pi} \Im \ln \det \langle \phi_i | e^{i \frac{2\pi}{L} \hat{x}} | \phi_j \rangle. \tag{41}$$

The electronic contribution to the polarization, $\vec{M}_{el}$, is then defined from the expectation value of x in the limit of a large system:

$$\vec{M}_{el} = \lim_{L \to \infty} \frac{e}{2\pi} \Im \ln \langle \Psi | e^{i \frac{2\pi}{L} x} | \Psi \rangle. \tag{42}$$

AIMD is very advantageous for the study of vibrational spectra. It explicitly accounts for the molecular environment of the solute, and its fluctuations, in the condensed phase and at finite temperature. It shares this property with force-field molecular dynamics. However, the AIMD potential energy surface is naturally anharmonic and frequency shifts induced by interaction with the solvent are well described. This anharmonicity however makes AIMD results hard at first to interpret in terms of vibrational modes. Recently, methods have been designed precisely to do so[110–112], in particular through the definition of effective or generalized normal mode coordinates[113–115].

Moreover, the electronic polarization discussed above is extracted from the electronic structure calculation. This is highly crucial[116] as illustrated on Figure 1 that compares the IR spectra of water obtained from the same dynamics but calculating the polarization in two ways. On the left, the polarization is obtained by a simple model assigning partial charges to the atoms (scaled to conserve the overall IR intensity), as in force-field calculations, neglecting water polarizability. On the right, it is calculated using the full ab initio polarization. We can see that the intensities of the signals are very different, the one obtained from AIMD being much closer to experiment. Comparison with experiment[117] shows a red-shift of the OH-stretch typical of DFT calculations using the Becke Lee Yang and Parr exchange-correlation functional[118,119]. For some peaks, such as the low frequency librational band at 60 cm^{-1} it is absent if polarizability is not taken into account. This is then crucial to interpret terahertz spectroscopy, which has emerged as a powerful tool to study solvation[120].

From the knowledge of the structure at the molecular level, knowledge of different spectroscopic signals, one can try to infer the main features of the interaction of the solute with water. To help drawing such a picture, AIMD has the further advantage that it gives direct access to the electronic structure. Figure 2, illustrate the electronic density for a system of 32 water molecules. As in standard molecular dynamics, the condensed phase is mimicked with periodic boundary conditions such that there is no surface for the system. In this figure, the electronic density can be seen to be delocalized and links the water molecules through the hydrogen bond network. Kohn-Sham orbitals, are similarly "molecular" orbitals that can span, and often do, the whole system. This is not how we would be tempted to represent the liquid, as a collection of molecules that can be described each by their molecular orbitals.

This molecular picture can be recovered by using maximally localized Wannier orbitals[121]. Any unitary transformation of the Kohn-Sham orbitals $\phi \to \tilde{\phi} = U^T \phi$ leaves

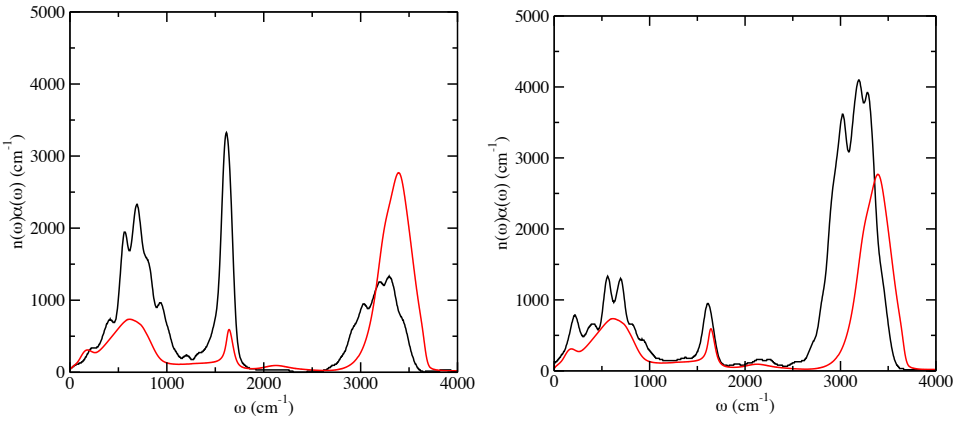


Figure 1: Calculated infrared spectra (black) and experimental infrared spectrum of water[117] (red) at $T = 300$ K. Left: the IR spectrum is calculated assigning a partial charge $q_O = 1.04\ e$ to oxygens and $q_H = 0.52\ e$ to the hydrogen atoms. Right: Full AIMD computation of the infrared spectrum. The calculation was performed with a 32 water molecules systems using the Car-Parrinello technique and Becke Lee Yang and Parr exchange-correlation functional[118,119]. The cutoff for the plane-wave basis set was set to 70 Ry.

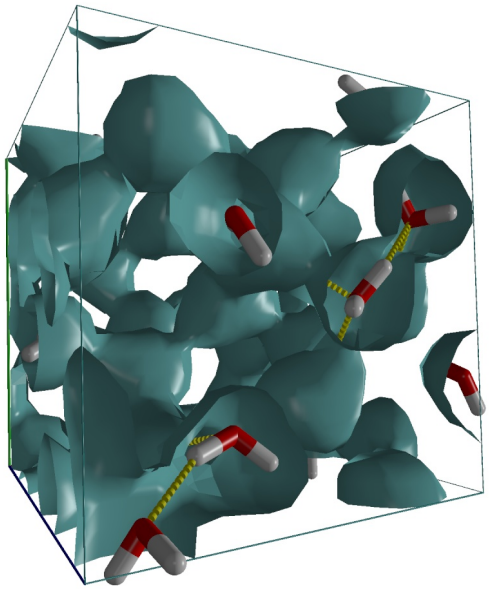


Figure 2: Snapshot of a 32 water moelcule AIMD simulation with periodic boundary conditions. The gery surface is an isocontour of the electronic density of the entire system.

both the electronic density $n(\mathbf{r})$ and the total energy invariant. Any unitary transformation of the Kohn-Sham orbitals is thus a valid set of orbitals. Canonical orbitals are a special set of such orbitals which diagonalize the Kohn-Sham Hamiltonian. Localized orbitals on

the other hand are obtained by finding the unitary transformation U so as to optimize the expectation value of a *two electrons* operator Ω:

$$\langle\Omega\rangle = \sum_{i=1}^{N} \langle\phi_i(1)\phi_i(2)|\Omega|\phi_i(1)\phi_i(2)\rangle \tag{43}$$

that describes the localization of the orbitals. In periodic boundary conditions, decribing localization is difficult because of the impossibility to define the position absolutely. A solution to this problem is to minimise

$$\Omega = \frac{-1}{(2\pi)^2} w_\alpha \ln|z_\alpha|^2 \; ; \; z_\alpha = \langle\Psi|e^{i\frac{2\pi}{L}r_\alpha}|\Psi\rangle, \tag{44}$$

w_α are weights assigned to each axis x, y and z. Maximally localized Wannier orbitals are defined from the choice of the unitary transformation which minimizes the sum of spreads, eq. (44), of the orbitals.

These localized orbitals have a well-defined center, referred to as the Wannier center in PBC and defined as[122–124]

$$r_\alpha^w = \frac{L}{2\pi} \Im \ln\langle\phi^w|e^{i\frac{2\pi}{L}r_\alpha}|\phi^w\rangle, \tag{45}$$

for each Wannier orbital ϕ^W. As can be seen from this definition the Wannier center is defined modulo the periodic boundary conditions.

Figure 3, shows the Wannier centers for two water molecules in liquid water, as well as two isocontours of their molecular density, obtained by summing the electronic densities of all maximally localized Wannier orbitals that can be assigned to a given molecule. The assignment of MLWO to water molecules is not ambiguous for this system, for example using a distance criteria. Their are four MLWO per water molecule, two representing the OH covalent bond and two the oxygen lone pairs. For non-metallic systems, the MLWO have an exponential decay at long range and are thus well localized[125,126]. It can be seen from Figure 3 that there is a very small overlap between the reconstructed molecular densities of two adjacent molecules, recovering a molecular description of the liquid.

This molecular decomposition of the total electronic density can be used for many purposes. They can serve for the calculation of molecular properties such as molecular dipole moments. The increase of the water dipole moments from 1.8 D in the gas phase to about 3 D in condensed phase is a signature of polarization induced by the hydrogen bond network[124]. In the case of ice, the molecular dipole moment can be related to the dielectric constant which is found to be in close agreement with experiments. Other molecular properties computed from the MLWO include polarisability[127], molecular energy levels[128], collision cross-sections[129] etc.

By identifying the main ingredients of the interaction between molecules it becomes possible to extend AIMD by developing force-field that include these features. To this respect also, the MLWO can also be used to construct force-field[130,131]. It allows for the evaluation of partial charges for atoms in the molecule, but also the long-range dispersion term and short range repulsion. Force-fields for alkali and halogen ions, reliable for describing the strucutre and the transport properties, where constructed in this way[132].

This is a further advantage of AIMD that should not be overlooked: the electronic state is not simply an intermediate for the computation of the nuclear forces but is an integrate

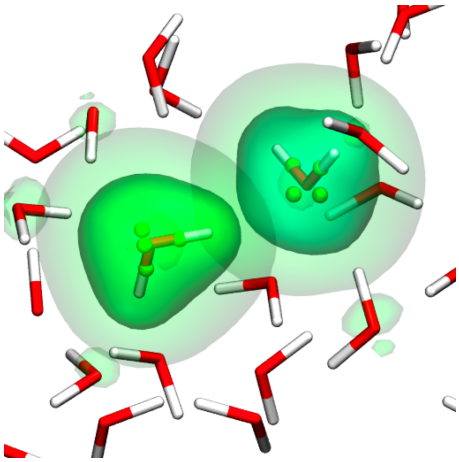


Figure 3: Wannier centers (green dots) of two water molecules in a 32 water molecules system with periodic boundary conditions. The green surfaces are isosurfaces of the reconstructed molecular electronic density from the Wannier orbitals assigned to each water molecule. The inner surfaces enclose 95% of the valence electron of each molecule, and the outer contour 99%.

part of the description of the system from which observables can be obtained and models of the interaction constructed.

3.3 Simulation of chemical reactions – free energy evaluation

For chemical reactions, AIMD allow to describe these situations where the solvent plays an active role, which turns out to be nearly the majority of cases. Examples of chemical reactions studied with AIMD include proton transfer reactions[56,133,134], proton transport in solvents[135,136], elimination and dissociation reactions[137–141] etc.[142–149,83,78,134,80,150] There too, AIMD allow for the description of extreme conditions and environment. Very recently, chemical reactions under intense electric fields were studied[151,152], reproducing in silico the Miller experiment showing the formation of organic molecules, that may be at the origin of life, from inorganic molecules under the influence of electric fields[152]. The mechanical activation of chemical reactions has also been studied with AIMD[153].

Because of the high computational cost of AIMD, only short timescales can be simulated, which renders the study of slow reactions difficult. However, the impressive development of acceleration techniques has helped to bridge these timescales. Because AIMD is a molecular dynamics on a particular potential energy surface, all techniques developed for molecular dynamics are readily available for AIMD, starting from the simple umbrella sampling, to blue moon sampling[154], transition path sampling[155], accelerated molecular dynamics[156,157], and metadynamics[158–160]. These methods, except transition path sampling, rely on the definition of a set of order parameters, which we will denote as s_α. These order parameters can distinguish the reactants from the products and acceleration techniques consist in forcing the system to sample the whole space of s_α and reconstruct the free energy $F(s_\alpha)$ on the order parameter map, from which the reaction mechanism can be extracted.

Many order parameters have been suggested and are available in codes such as PLUMED[161]. They depend on the system configuration. It can be an interatomic distance, a coordination number or more involved collective variable describing the topology of the structure[162]. However, at difference with force-field based molecular dynamics, it has been suggested that the order parameters s_α can directly involve the electrons, such that the HOMO-LUMO gap[163,164].

Also specific to AIMD, methods have been developed for the calculation of acidity constants[145,149,165,166] or oxydoreduction potentials of ions or small molecules[167,83,78,168,79,80,165,169,166]. While the acidity constant can be computed as the free energy of dissociation reaction, taking as order parameter the distance to the excess proton from the acid site, it can also be computed as the free energy to insert a proton at the acid site. Oxydation or reduction free energies are similarly obtained as the free energy to extract or insert an electron in the system. Both acidity constant and oxydoreduction potentials are then evaluated in a consistent way through thermodynamical integration. The free energy difference $\Delta F_{1\to 2} = F(S_2) - F(S_1)$ between two states of the system, S_1 and S_2, which are described by two potential energy surfaces V_1 and V_2, can be written as

$$\Delta F_{1\to 2} = \int_0^1 \langle V_2 - V_1 \rangle_\lambda \, \mathrm{d}\lambda, \tag{46}$$

where $\langle \cdot \rangle_\lambda$ denotes an average at a given λ interpolating between S_1 and S_2, average of the canonical distribution at temperature T with a potential energy surface $\lambda V_2 + (1-\lambda)V_1$. For example S_1 and V_1 can describe a system with N electron (oxydized) while S_2 and V_2 describe the same system with $N+1$ electrons (reduced). These methods are adapted to the computaion of the reaction free energy of proton coupled electron transfer. In the general case, the main difficulty resides in the fact that a charged particle is inserted in the system and absolute energies are difficult to obtain because of the periodic boundary conditions that make the system surface disappear[170–172].

4 Extensions

We have described the main flavour of ab initio simulations, ab initio molecular dynamics, that relies on the Born-Oppenheimer separation and a classical representation of the nuclei motion. Extension have been proposed however.

Path-integral methods have been introduced to have a quantum representation of the nuclei[173,174]. In these methods, the quantum mechanical partition function is mapped onto the classical partition function of necklaces, with one necklace representing on nuclei. The beads of these necklaces evolve under the influence of the other beads in the same necklace and on the potential energy surface[175]. In ab initio path-integral method, this potential energy contribution is obtained from an electronic structure calculation. This is particularly of interest for proton transfer reactions and transport where quantum effects are important because of the light mass of the proton[176,177,56,178,179]. Recently, the proton quantum momentum distribution has been obtained in water[180–182] and ice[183] and these calculation show that a proper description of the anharmonicity of the OH stretch of water, linked to the hydrogen bond, is crucial, and is only described correctly in ab initio approaches.

There are also situations where the hypothesis that has been made that the system remains in the electronic ground state is not valid. This is typically the case upon pho-

toexcitation. Upon photoexcitation, the system is put in an electronic excited state and it relaxes through non-adiabatic transitions between the adiabatic states. There are ab initio techniques developed to simulate these processes, with explicit electronic structure calculations of the adiabatic states and the non-adiabatic couplings between them[184–190]. Other approaches employ an explicit time dependent propagation of the electronic state, so-called Ehrenfest dynamics that couples in a mean field the electrons and nuclei, for very non-adiabatic and highly energetic situation as in the case of ionization induced by irradiation[191,192].

Finally, DFT has its limitations. We discussed for example the lack of long range dispersion. Given the computer power now available and the advent of new algorithms, it is now possible to perform ab initio simulations with wavefunction based methods such as MP2[52]. It has been shown that in the case of water, the correct density can then be obtained at this level of electronic structure theory, at ambient pressure and temperature.

5 Conclusions

AIMD is not just an on the fly calculation of forces using an electronic structure calculation. The data of the electronic state along the nuclear trajectory is very valuable to get insight in the interactions at play in the system. Moreover, it pushes the development of prediction of spectroscopic measurement that aim at directly probing the microscopic structure and the nature of the interactions. It is also of great value to have a direct comparison with experiments. This helps in extracting the necessary components to build models upon the high level of simulation that AIMD represents. In parallel, free energy calculations are now accessible, in particular for the study of chemical reactions. Through these, the search for reaction mechanism has gone from the observation of a few trajectories to a meaningful statistical analysis of the free energy surface.

AIMD has then become a very powerful and reliable technique; it is a very active and changing field, which makes it fascinating. We have given here a flavour of the techniques involved and of the type of system studied. However, its field of application is very large, which is why the method is so attractive, and only a fraction of it has been explored. No doubt that there are new discoveries to come from using AIMD.

Acknowledgments

I am obviously indebted to a number of collaborators in the field and developers of the AIMD codes, some of which are now good friends. In particular I would like to warmly thank Michiel Sprik who introduced me in the field of Car-Parrinello molecular dynamics. I would also like to thank Dominik Marx for inviting me to participate to this tutorial and the invitation to write these lecture notes.

References

1. R. Car and M. Parrinello, *Unified Approach for Molecular Dynamics and Density-Functional Theory*, Phys. Rev. Lett., **55**, 2471, 1985.

2. D. Marx and J. Hutter, "Ab initio molecular dynamics: Theory and implementation; modern methods and algorithms of quantum chemistry", in: NIC Series, J. Grotendorst, (Ed.), vol. 1. Forschungszentrum Jülich, 2000.
3. J. Hutter, "Large scale density functional calculations", in: Multiscale Modelling and Simulation, S. Attinger and P. Koumoutsakos, (Eds.), Lecture Notes in Computational Science and Engineering, p. 39. Springer, Heidelberg, 2004.
4. Dominik Marx and Jürg Hutter, *Ab initio Molecular Dynamics – Basic Theory and Advanced Methods*, Cambridge University Press, 2012.
5. R. Vuilleumier, "Density Functional Theory Based Ab Initio Molecular Dynamics Using the Car-Parrinello Approach", in: Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1, Mauro Ferrario Professor, Giovanni Ciccotti Professor, and Kurt Binder Professor, (Eds.), number 703 in Lecture Notes in Physics, pp. 223–285. Springer Berlin Heidelberg, 2006.
6. Rodolphe Vuilleumier and Michiel Sprik, *Electronic properties of hard and soft ions in solution: Aqueous Na^+ and Ag^+ compared*, J. Chem. Phys., **115**, no. 8, 3454–3468, 2001.
7. Christophe Gourlaouen, Hélène Gérard, Jean-Philip Piquemal, and Olivier Parisel, *Understanding Lead Chemistry from Topological Insights: The Transition between Holo- and Hemidirected Structures within the [Pb(CO)n]2+ Model Series*, Chemistry ? A European Journal, **14**, no. 9, 2730–2743, Mar. 2008.
8. M. E. Tuckerman, *Ab initio molecular dynamics: Basic concepts, current trends and novel applications*, J. Phys. Condens. Matt., **14**, R1297, 2002.
9. M. Born and J. R. Oppenheimer, Annalen der Physik, **84**, 457–484, 1927.
10. Max Born and Kun Huang, *Dynamical Theory of Crystal Lattices.*, Oxford University Press, Oxford, 1954.
11. D. M. Ceperley and B. J. Alder, *Ground State of the Electron Gas by a Stochastic Method*, Phys. Rev. Lett., **45**, 566, 1980.
12. C. Pierleoni, D. M. Ceperley, and M. Holzmann, *Coupled Electron-Ion Monte Carlo Calculations of Dense Metallic Hydrogen*, Phys. Rev. Lett., **93**, 146402, 2004.
13. R. M. Dreizler and E. K. U. Gross, *Density Functional Theory*, Springer-Verlag, Berlin Heidelberg, 1990.
14. Robert G. Parr and Yang Weitao, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, 1994.
15. Richard M. Martin, *Electronic Structure : Basic Theory and Practical Methods*, Cambridge University Press, 2004.
16. P. Hohenberg and W. Kohn, *Inhomogeneous Electron Gas*, Phys. Rev., **136**, B864–B871, 1964.
17. Mel Levy, *Universal Variational Functionals of Electron Densities, First-Order Density Matrices, and Natural Spin-Orbitals and Solution of the v-Representability Problem*, Proc. Nat. Acad. Science, **76**, 6062–6065, 1979.
18. W. Kohn and L. J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects*, Phys. Rev., **140**, A1133–A1138, 1965.
19. Kieron Burke, *Perspective on density functional theory*, The Journal of Chemical Physics, **136**, no. 15, 150901–150901–9, Apr. 2012.
20. Axel D. Becke, *Perspective: Fifty years of density-functional theory in chemical physics*, The Journal of Chemical Physics, **140**, no. 18, 18A301, May 2014.

21. Kieron Burke and Lucas O. Wagner, *DFT in a nutshell*, International Journal of Quantum Chemistry, **113**, no. 2, 96–101, Jan. 2013.
22. Dmitrij Rappoport, Nathan R. M. Crawford, Filipp Furche, and Kieron Burke, "Approximate Density Functionals: Which Should I Choose?", in: Encyclopedia of Inorganic Chemistry. John Wiley & Sons, Ltd, 2006.
23. M. P. Teter, M. C. Payne, and D. C. Allen, *Solution of Schrödinger's equation for large systems*, Phys. Rev. B, **40**, 12255, 1989.
24. D. R. Hamann, M. Schlüter, and C. Chiang, *Norm-Conserving Pseudopotentials*, Phys. Rev. Lett., **43**, 1494, 1979.
25. G. B. Bachelet, D. R. Hamann, and M. Schlüter, *Pseudopotentials that work: From H to Pu*, Phys. Rev. B, **26**, 4199, 1982.
26. N. Troullier and J. L. Martins, *Efficient pseudopotentials for plane-wave calculations*, Phys. Rev. B, **43**, 1993, 1991.
27. S. Goedecker, M. Teter, and J. Hutter, *Seperable Dual-Space Gaussian Pseudopotentials*, Phys. Rev. B, **54**, 1703–1710, 1996.
28. C. Hartwigsen, S. Goedecker, and J. Hutter, *Relativistic separable dual-space Gaussian pseudopotentials from H to Rn*, Phys. Rev. B, **58**, 3641–3662, 1998.
29. D. Vanderbilt, *Optimally smooth norm-conserving pseudopotentials*, Phys. Rev. B, **32**, 8412–8415, 1985.
30. D. Vanderbilt, *Soft self-consistent pseudopotentials in a generalized eigenvalue formalism*, Phys. Rev. B, **41**, 7892–7895, 1990.
31. K. Laasonen, R. Car, C. Lee, and D. Vanderbilt, *Implementation of ultrasoft pseudopotentials in ab initio molecular dynamics*, Phys. Rev. B, **43**, 6796–6799, 1991.
32. K. Laasonen, A. Pasquarello, R. Car, C. Lee, and D. Vanderbilt, *Car-Parrinello molecular dynamics with Vanderbilt ultrasoft pseudopotentials*, Phys. Rev. B, **47**, 10142–10153, 1993.
33. J. Hutter, M. E. Tuckerman, and M. Parrinello, *Integrating the Car-Parrinello equations III*, J. Chem. Phys., **102**, 859, 1995.
34. J. Vandevondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, *Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach*, Comp. Phys. Comm., **167**, 103–128, 2005.
35. José M. Soler, Emilio Artacho, Julian D. Gale, Alberto García, Javier Junquera, Pablo Ordejón, and Daniel Sánchez-Portal, *The SIESTA method for ab initio order-N materials simulation*, Journal of Physics: Condensed Matter, **14**, no. 11, 2745, Mar. 2002.
36. "Cpmd", Copyright IBM Corp 1990-2001, Copyright MPI für Festkörperforschung Stuttgart, 1997-2004.
37. Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L. Chiarotti, Matteo Cococcioni, Ismaila Dabo, Andrea Dal Corso, Stefano de Gironcoli, Stefano Fabris, Guido Fratesi, Ralph Gebauer, Uwe Gerstmann, Christos Gougoussis, Anton Kokalj, Michele Lazzeri, Layla Martin-Samos, Nicola Marzari, Francesco Mauri, Riccardo Mazzarello, Stefano Paolini, Alfredo Pasquarello, Lorenzo Paulatto, Carlo Sbraccia, Sandro Scandolo, Gabriele Sclauzero, Ari P. Seitsonen, Alexander Smogunov, Paolo Umari, and Renata M. Wentzcovitch, *QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials*, Journal of Physics: Condensed

Matter, **21**, no. 39, 395502, Sept. 2009.
38. J. VandeVondele and J. Hutter, *An efficient orbital transformation method for electronic structure calculations*, J. Chem. Phys., **118**, 4365–4369, 2003.
39. Jirí Kolafa, *Time-reversible always stable predictor–corrector method for molecular dynamics of polarizable molecules*, Journal of Computational Chemistry, **25**, no. 3, 335–342, Feb. 2004.
40. Thomas D. Kühne, Matthias Krack, Fawzi R. Mohamed, and Michele Parrinello, *Efficient and Accurate Car-Parrinello-like Approach to Born-Oppenheimer Molecular Dynamics*, Physical Review Letters, **98**, no. 6, 066401, Feb. 2007.
41. I.-F. W. Kuo, C. J. Mundy, M. J. McGrath, J. I. Siepmann, J. VandeVondele, M. Sprik, J. Hutter, B. Chen, M. L. Klein, F. Mohamed, M. Krack, and M. Parrinello, *Liquid water from first principles, investigation of different sampling approaches*, J. Phys. Chem. B, **108**, 12990, 2004.
42. M.J. McGrath, J.I. Siepmann, I.-F. W. Kuo, C. J. Mundy, J. VandeVondele, J. Hutter, F. Mohamed, and M. Krack, *First Principles: Application to Liquid Water at Ambient Conditions*, ChemPhysChem, **6**, 1894–1901, 2005.
43. M.J. McGrath, J.I. Siepmann, I.-F. W. Kuo, C. J. Mundy, J. VandeVondele, M. Sprik, J. Hutter, F. Mohamed, M. Krack, and M. Parrinello, *Toward a Monte Carlo program for simulating vapor-liquid phase equilibria from first principles*, Comp. Phys. Comm., **169**, 289–294, 2005.
44. Soohaeng Yoo and Sotiris S. Xantheas, *Communication: The effect of dispersion corrections on the melting temperature of liquid water*, The Journal of Chemical Physics, **134**, no. 12, 121105, 2011.
45. Jochen Schmidt, Joost VandeVondele, I.-F. William Kuo, Daniel Sebastiani, J. Ilja Siepmann, Jürg Hutter, and Christopher J. Mundy, *Isobaric–Isothermal Molecular Dynamics Simulations Utilizing Density Functional Theory: An Assessment of the Structure and Density of Water at Near-Ambient Conditions*, The Journal of Physical Chemistry B, **113**, no. 35, 11959–11964, Sept. 2009.
46. M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, *Van der Waals Density Functional for General Geometries*, Physical Review Letters, **92**, no. 24, 246401, June 2004.
47. Stefan Grimme, *Semiempirical GGA-type density functional constructed with a long-range dispersion correction*, Journal of Computational Chemistry, **27**, no. 15, 1787–1799, Nov. 2006.
48. Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg, *A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu*, The Journal of Chemical Physics, **132**, no. 15, 154104, Apr. 2010.
49. I-Chun Lin, Ari P. Seitsonen, Maurício D. Coutinho-Neto, Ivano Tavernelli, and Ursula Rothlisberger, *Importance of van der Waals Interactions in Liquid Water*, The Journal of Physical Chemistry B, **113**, no. 4, 1127–1131, Jan. 2009.
50. Romain Jonchiere, Ari P. Seitsonen, Guillaume Ferlat, A. Marco Saitta, and Rodolphe Vuilleumier, *Van der Waals effects in ab initio water at ambient and supercritical conditions*, Journal of Chemical Physics, **135**, no. 15, 154503, Oct. 2011.
51. I-Chun Lin, Ari P. Seitsonen, Ivano Tavernelli, and Ursula Rothlisberger, *Structure and Dynamics of Liquid Water from ab Initio Molecular Dynamics: Comparison of BLYP, PBE, and revPBE Density Functionals with and without van der Waals Corrections*, Journal of Chemical Theory and Computation, **8**, no. 10, 3902–3910, Oct.

2012.
52. Mauro Del Ben, Mandes Schönherr, Jürg Hutter, and Joost VandeVondele, *Bulk Liquid Water at Ambient Temperature and Pressure from MP2 Theory*, The Journal of Physical Chemistry Letters, **4**, no. 21, 3753–3759, Nov. 2013.
53. D. Marx, J. Hutter, and M. Parrinello, *Density functional study of small aqueous Be2+ clusters*, Chem. Phys. Lett., **241**, 457, 1995.
54. M. E. Tuckerman, K. Laasonen, M. Sprik, and M. Parrinello, *Ab initio molecular dynamics simulation of the solvation and transport of H3O+ and OH- ions in water*, J. Phys. Chem., **99**, 5749, 1995.
55. D. Marx, M. Sprik, and M. Parrinello, *Ab initio molecular dynamics of ion solvation. The case of Be2+ in water*, Chem. Phys. Lett., **273**, 360–366, 1997.
56. D. Marx, M.E. Tuckerman, J. Hutter, and M. Parrinello, *The nature of the hydrated excess proton in water*, Nature, **397**, 601–604, 1999.
57. M. Tuckerman, K. Laasonen, M. Sprik, and M. Parrinello, *Ab initio molecular dynamics simulation of the solvation and transport of hydronium and hydroxyl ions in water*, J. Chem. Phys., **103**, no. 1, 150–161, 1995.
58. L. M. Ramaniah, M. Bernasconi, and M. Parrinello, *Ab initio molecular-dynamics simulation of K+ solvation in water*, J. Chem. Phys., **111**, no. 4, 1587–1591, 1999.
59. C. J. Mundy, J. Hutter, and M. Parrinello, *Microsolvation and chemical reactivity of sodium and water clusters*, J. Am. Chem. Soc, **122**, 4837–4838, 2000.
60. Zhi-Pan Liu, P. Hu, and Ali Alavi, *Mechanism for the high reactivity of CO oxidation on a ruthenium-oxide*, J. Chem. Phys., **114**, 5956, 2001.
61. Titus S. van Erp and Evert Jan Meijer, *Ab initio molecular dynamics study of aqueous solvation of ethanol and ethylene*, J. Chem. Phys., **118**, no. 19, 8831–8840, 2003.
62. J. M. Heuft and E. J. Meijer, *Density functional theory based molecular-dynamics study of aqueous chloride solvation*, J. Chem. Phys., **119**, no. 22, 11788–11791, 2003.
63. M. E. Tuckerman, D. Marx, and M. Parrinello, *The nature and transport mechanism of hydrated hydroxide ions in aqueous solution*, Nature, **417**, 925–929, 2002.
64. B. Kirchner, J. Stubbs, and D. Marx, *Fast Anomalous Diffusion of Small Hydrophobic Species in Water*, Phys. Rev. Lett., **89**, 215901, 2002.
65. B. Kirchner and J. Hutter, *The structure of a DMSO/Water mixture from Car-Parrinello simulations*, Chem. Phys. Lett., **364**, 497–502, 2002.
66. Takashi Ikeda, Masaru Hirata, and Takaumi Kimura, *Hydration of Y^{3+} ion: A Car-Parrinello molecular dynamics study*, J. Chem. Phys., **122**, no. 2, 024510, 2005.
67. Sergei Izvekov and Gregory A. Voth, *Ab initio molecular-dynamics simulation of aqueous proton solvation and transport revisited*, J. Chem. Phys., **123**, no. 4, 044505, 2005.
68. M. Boero, M. Parrinello, K. Terakura, T. Ikeshoji, and C. C. Liew, *First-Principles Molecular-Dynamics Simulations of a Hydrated Electron in Normal and Supercritical Water*, Phys. Rev. Lett., **90**, 226403, 2003.
69. C. Nicolas, R. Spezia, A. Boutin, and R. Vuilleumier, *Molecular Dynamics simulations of a silver atom in water: dipolar excitonic state evidence*, Phys. Rev. Lett., **91**, 208304, 2003.
70. M.-P. Gaigeot and M. Sprik, *Ab initio molecular dynamics study of uracil in aqueous solution*, J. Phys. Chem. B, **108**, 7458, 2004.
71. L. Bernasconi, M. Sprik, and J. Hutter, *Hartree-Fock exchange in time dependent density functional theory: Application to charge transfer excitations in solvated molecu-*

lar systems, Chem. Phys. Lett., **394**, 141, 2004.
72. F. C. Lightstone, E. Schwegler, M. Allesch, F. Gygi, and G. Galli, *A first-principles molecular dynamics study of calcium in water*, ChemPhysChem, **6**, 1745–1749, 2005.
73. J. M. Heuft and E. J. Meijer, *Density functional theory based molecular-dynamics study of aqueous fluoride solvation*, J. Chem. Phys., **122**, no. 9, 094501, 2005.
74. J. M. Heuft and E. J. Meijer, *Density functional theory based molecular-dynamics study of aqueous iodide solvation*, J. Chem. Phys., **123**, no. 9, 094506, 2005.
75. J. VandeVondele and M. Sprik, *A molecular dynamics study of the hydroxyl radical in solution applying self-interaction corrected density functional methods*, Phys. Chem. Chem. Phys., **7**, 1363, 2005.
76. P. Hunt and M. Sprik, *On the position of the highest molecular orbital in aqueous solutions of simple ions*, ChemPhysChem, **6**, 1805, 2005.
77. Elske J. M. Leenders, Peter G. Bolhuis, and Evert Jan Meijer, *Microscopic Picture of the Aqueous Solvation of Glutamic Acid*, Journal of Chemical Theory and Computation, **4**, no. 6, 898–907, June 2008.
78. J. Blumberger and M. Sprik, *Free energy of oxidation of metal aqua Ions by an enforced change of Coordination*, J. Phys. Chem. B., **108**, 6529, 2004.
79. J. Blumberger and M. Sprik, *Ab initio molecular dynamics simulation of the aqueous Ru2+/Ru3+ redox reaction: The Marcus perspective*, J. Phys. Chem. B, **109**, 6793, 2005.
80. Yoshitaka Tateyama, Jochen Blumberger, Michiel Sprik, and Ivano Tavernelli, *Density-functional molecular-dynamics study of the redox reactions of two anionic, aqueous transition-metal complexes*, J. Chem. Phys., **122**, no. 23, 234505, 2005.
81. Xiandong Liu, Evert Jan Meijer, Xiancai Lu, and Rucheng Wang, *Ab initio Molecular Dynamics Study of Fe-Containing Smectites*, Clays and Clay Minerals, **58**, no. 1, 89–96, Jan. 2010.
82. Leonardo Bernasconi, Jochen Blumberger, Michiel Sprik, and Rodolphe Vuilleumier, *Density functional calculation of the electronic absorption spectrum of Cu^+ and Ag^+ aqua ions*, J. Chem. Phys., **121**, no. 23, 11885–11899, 2004.
83. J. Blumberger, L. Bernasconi, I. Tavernelli, R. Vuilleumier, and M. Sprik, *Electronic structure and solvation of copper and silver Ions: A theoretical picture of a model aqueous redox reaction*, J. Am. Chem. Soc., **126**, 3928, 2004.
84. Xiandong Liu, Xiancai Lu, Rucheng Wang, Huiqun Zhou, and Shijin Xu, *Speciation of gold in hydrosulphide-rich ore-forming fluids: Insights from first-principles molecular dynamics simulations*, Geochimica et Cosmochimica Acta, **75**, no. 1, 185–194, Jan. 2011.
85. Gleb S. Pokrovski, Jacques Roux, Guillaume Ferlat, Romain Jonchiere, Ari P. Seitsonen, Rodolphe Vuilleumier, and Jean-Louis Hazemann, *Silver in geological fluids from in situ X-ray absorption spectroscopy and first-principles molecular dynamics*, Geochimica et Cosmochimica Acta, **106**, 501–523, Apr. 2013.
86. Yuan Mei, David M. Sherman, Weihua Liu, and Joël Brugger, *Complexation of gold in S_3^--rich hydrothermal fluids: Evidence from ab-initio molecular dynamics simulations*, Chemical Geology, **347**, 34–42, June 2013.
87. Yuan Mei, David M. Sherman, Weihua Liu, and Joël Brugger, *Ab initio molecular dynamics simulation and free energy exploration of copper(I) complexation by chloride and bisulfide in hydrothermal fluids*, Geochimica et Cosmochimica Acta, **102**, 45–64,

Feb. 2013.
88. Yannick Jeanvoine, Pere Miró, Fausto Martelli, Christopher J. Cramer, and Riccardo Spezia, *Electronic structure and bonding of lanthanoid(III) carbonates*, Physical Chemistry Chemical Physics, **14**, no. 43, 14822–14831, Oct. 2012.
89. Fausto Martelli, Yannick Jeanvoine, Thomas Vercouter, César Beuchat, Rodolphe Vuilleumier, and Riccardo Spezia, *Hydration properties of lanthanoid(III) carbonate complexes in liquid water determined by polarizable molecular dynamics simulations*, Physical Chemistry Chemical Physics, **16**, no. 8, 3693–3705, Jan. 2014.
90. Riccardo Spezia, Cesar Beuchat, Rodolphe Vuilleumier, Paola D?Angelo, and Laura Gagliardi, *Unravelling the Hydration Structure of ThX4 (X = Br, Cl) Water Solutions by Molecular Dynamics Simulations and X-ray Absorption Spectroscopy*, The Journal of Physical Chemistry B, **116**, no. 22, 6465–6475, June 2012.
91. Xiandong Liu, Xiancai Lu, Rucheng Wang, and Evert Jan Meijer, *Understanding hydration of Zn2+ in hydrothermal fluids with ab initio molecular dynamics*, Physical Chemistry Chemical Physics, **13**, no. 29, 13305–13309, July 2011.
92. Carole Bresson, Riccardo Spezia, Stéphane Esnouf, Pier Lorenzo Solari, Stéphanie Coantic, and Christophe Den Auwer, *A combined spectroscopic and theoretical approach to investigate structural properties of Co(ii)/Co(iii) tris-cysteinato complexes in aqueous medium*, New Journal of Chemistry, **31**, no. 10, 1789, 2007.
93. Riccardo Spezia, Carole Bresson, Christophe Den Auwer, and Marie-Pierre Gaigeot, *Solvation of Co(III)-Cysteinato Complexes in Water: A DFT-based Molecular Dynamics Study*, The Journal of Physical Chemistry B, **112**, no. 20, 6490–6499, May 2008.
94. David Prendergast and Giulia Galli, *X-Ray Absorption Spectra of Water from First Principles Calculations*, Physical Review Letters, **96**, no. 21, 215502, May 2006.
95. Marcella Iannuzzi and J?rg Hutter, *Inner-shell spectroscopy by the Gaussian and augmented plane wave method*, Physical Chemistry Chemical Physics, **9**, no. 13, 1599, 2007.
96. Sittipong Komin, Christian Gossens, Ivano Tavernelli, Ursula Rothlisberger, and Daniel Sebastiani, *NMR Solvent Shifts of Adenine in Aqueous Solution from Hybrid QM/MM Molecular Dynamics Simulations*, The Journal of Physical Chemistry B, **111**, no. 19, 5225–5232, May 2007.
97. Ute F. Röhrig and Daniel Sebastiani, *NMR Chemical Shifts of the Rhodopsin Chromophore in the Dark State and in Bathorhodopsin: A Hybrid QM/MM Molecular Dynamics Study*, The Journal of Physical Chemistry B, **112**, no. 4, 1267–1274, Jan. 2008.
98. Xiang Yang Guo, Tobias Watermann, Shane Keane, Christoph Allolio, and Daniel Sebastiani, *First Principles Calculations of NMR Chemical Shifts of Liquid Water at an Amorphous Silica Interface*, Zeitschrift für Physikalische Chemie International journal of research in physical chemistry and chemical physics, **226**, no. 11-12, 1415–1424, 2012.
99. Hossam Elgabarty, Milian Wolff, Adrian Glaubitz, Dariush Hinderberger, and Daniel Sebastiani, *First principles calculation of inhomogeneous broadening in solid-state cw-EPR spectroscopy*, Physical Chemistry Chemical Physics, **15**, no. 38, 16082–16089, Sept. 2013.
100. Leonardo Bernasconi, Michiel Sprik, and Jurg Hutter, *Time dependent density func-*

tional theory study of charge-transfer and intramolecular electronic excitations in acetone–water systems, J. Chem. Phys., **119**, no. 23, 12417–12431, 2003.

101. L. Bernasconi and M. Sprik, *Time-dependent density functional theory description of on-site electron repulsion and ligand field effects in the optical spectrum of hexa-aquoruthenium(II) in solution*, J. Phys. Chem. B, **109**, 12222, 2005.
102. Robert Seidel, Manfred Faubel, Bernd Winter, and Jochen Blumberger, *Single-Ion Reorganization Free Energy of Aqueous Ru(bpy)32+/3+ and Ru(H2O)62+/3+ from Photoemission Spectroscopy and Density Functional Molecular Dynamics Simulation*, Journal of the American Chemical Society, **131**, no. 44, 16127–16137, Nov. 2009.
103. Robert Seidel, Stephan Thürmer, Jan Moens, Paul Geerlings, Jochen Blumberger, and Bernd Winter, *Valence Photoemission Spectra of Aqueous $Fe2^{+}/3^{+}$ and $[Fe(CN)6]4^{-}/3^{-}$ and Their Interpretation by DFT Calculations*, The Journal of Physical Chemistry B, **115**, no. 40, 11671–11677, Oct. 2011.
104. Pier Luigi Silvestrelli, M. Bernasconi, and Michele Parrinello, *Ab initio infrared spectrum of liquid water*, Chem. Phys. Lett., **277**, no. 5-6, 478–482, 1997.
105. R. Resta, *Quantum-Mechanical Position Operator in Extended Systems*, Phys. Rev. Lett., **80**, 1800–1803, 1998.
106. R. Resta and S. Sorella, *Electron Localization in the Insulating State*, Phys. Rev. Lett., **82**, 370–373, 1999.
107. R. Resta, *Macroscopic polarization in crystalline dielectrics: the geometric phase approach*, Rev. Mod. Phys., **66**, 1994, 899-915.
108. D. Vanderbilt and R. D. King-Smith, *Electric polarization as a bulk quantity and its relation to surface charge*, Phys. Rev. B, **48**, 4442–4455, 1993.
109. R. D. King-Smith and D. Vanderbilt, *Theory of polarization of crystalline solids*, Phys. Rev. B, **47**, 1651–1654, 1993.
110. Matthias Schmitz and Paul Tavan, *Vibrational spectra from atomic fluctuations in dynamics simulations. I. Theory, limitations, and a sample application*, The Journal of Chemical Physics, **121**, no. 24, 12233, 2004.
111. Matthias Schmitz and Paul Tavan, *Vibrational spectra from atomic fluctuations in dynamics simulations. II. Solvent-induced frequency fluctuations at femtosecond time resolution*, The Journal of Chemical Physics, **121**, no. 24, 12247, 2004.
112. Gerald Mathias, Sergei D. Ivanov, Alexander Witt, Marcel D. Baer, and Dominik Marx, *Infrared Spectroscopy of Fluxional Molecules from (ab Initio) Molecular Dynamics: Resolving Large-Amplitude Motion, Multiple Conformations, and Permutational Symmetries*, Journal of Chemical Theory and Computation, **8**, no. 1, 224–234, Jan. 2012.
113. M. Martinez, M.-P. Gaigeot, D. Borgis, and R. Vuilleumier, *Extracting effective normal modes from equilibrium dynamics at finite temperature*, The Journal of Chemical Physics, **125**, no. 14, 144106–144106–14, Oct. 2006.
114. Marie-Pierre Gaigeot, Michaël Martinez, and Rodolphe Vuilleumier, *Infrared spectroscopy in the gas and liquid phase from first principle molecular dynamics simulations: application to small peptides*, Molecular Physics, **105**, no. 19-22, 2857–2878, Oct. 2007.
115. Gerald Mathias and Marcel D. Baer, *Generalized Normal Coordinates for the Vibrational Analysis of Molecular Dynamics Simulations*, Journal of Chemical Theory and Computation, **7**, no. 7, 2028–2039, July 2011.
116. Rodolphe Vuilleumier, *Atomic partial charges in condensed phase from an exact sum*

rule for infrared absorption, Molecular Physics, **112**, no. 9-10, 1457–1462, May 2014.

117. J. E. Bertie and Z. Lan, *Infrared Intensities of Liquids XX: The Intensity of the OH Stretching Band of Liquid Water Revisited, and the Best Current Values of the Optical Constants of $H_2O(l)$ at 25°C*, App. Spec., **50**, 1047, 1996.
118. A. D. Becke, *Density-functional exchange-energy approximation with correct asymptotic behavior*, Phys. Rev. A, **38**, 3098–3100, 1988.
119. C. Lee, W. Yang, and R. G. Parr, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density*, Phys. Rev. B, **37**, 785–789, 1988.
120. Matthias Heyden, Jian Sun, Stefan Funkner, Gerald Mathias, Harald Forbert, Martina Havenith, and Dominik Marx, *Dissecting the THz spectrum of liquid water from first principles via correlations in time and space*, Proceedings of the National Academy of Sciences, **107**, no. 27, 12068–12073, June 2010.
121. Nicola Marzari, Arash A. Mostofi, Jonathan R. Yates, Ivo Souza, and David Vanderbilt, *Maximally localized Wannier functions: Theory and applications*, Reviews of Modern Physics, **84**, no. 4, 1419–1475, Oct. 2012.
122. P. L. Silvestrelli, *Maximally localized Wannier functions for simulations with supercells of general symmetry*, Phys. Rev. B, **59**, 9703–9706, 1999.
123. G. Berghold, C. J. Mundy, A. H. Romero, J. Hutter, and M. Parrinello, *General and Efficient Algorithms for Obtaining Maximally-Localized Wannier Functions*, Phys. Rev. B, **61**, 10040–10048, 2000.
124. P. L. Silvestrelli and M. Parrinello, *Water Molecule Dipole in the Gas and in the Liquid Phase*, Phys. Rev. Lett., **82**, 3308–3311, 1999.
125. Lixin He and David Vanderbilt, *Exponential Decay Properties of Wannier Functions and Related Quantities*, Physical Review Letters, **86**, no. 23, 5341–5344, June 2001.
126. Christian Brouder, Gianluca Panati, Matteo Calandra, Christophe Mourougane, and Nicola Marzari, *Exponential Localization of Wannier Functions in Insulators*, Physical Review Letters, **98**, no. 4, 046402, Jan. 2007.
127. M. Salanne, R. Vuilleumier, P. A Madden, C. Simon, P. Turq, and B. Guillot, *Polarizabilities of individual molecules and ions in liquids from first principles*, Journal of Physics: Condensed Matter, **20**, 494207, 2008.
128. P. Hunt, M. Sprik, and R. Vuilleumier, *Thermal versus electronic broadening in the density of states of liquid water*, Chem. Phys. Lett., **376**, 68, 2003.
129. M. L. de Sanctis, M.-F. Politis, R. Vuilleumier, C. Stia, and O. Fojón, *Triple differential cross sections for liquid water ionization by impact of fast electrons*, Journal of Physics: Conference Series, **488**, no. 5, 052024, Apr. 2014.
130. Benjamin Rotenberg, Mathieu Salanne, Christian Simon, and Rodolphe Vuilleumier, *From Localized Orbitals to Material Properties: Building Classical Force Fields for Nonmetallic Condensed Matter Systems*, Physical Review Letters, **104**, no. 13, 138301, Mar. 2010.
131. Mathieu Salanne, Benjamin Rotenberg, Sandro Jahn, Rodolphe Vuilleumier, Christian Simon, and Paul A. Madden, *Including many-body effects in models for ionic liquids*, Theoretical Chemistry Accounts, **131**, no. 3, Feb. 2012.
132. Sami Tazi, John J. Molina, Benjamin Rotenberg, Pierre Turq, Rodolphe Vuilleumier, and Mathieu Salanne, *A transferable ab initio based force field for aqueous ions*, The Journal of Chemical Physics, **136**, no. 11, 114507–114507–12, Mar. 2012.
133. Y. Liu and M. E. Tuckerman, *Protonic defects in hydrogen bonded liquids: Structure*

and dynamics in ammonia and comparison with water, J. Phys. Chem. B, **105**, 6598, 2001.
134. Phineus R. L. Markwick, Nikos L. Doltsinis, and Dominik Marx, *Targeted Car-Parrinello molecular dynamics: Elucidating double proton transfer in formic acid dimer*, J. Chem. Phys., **122**, no. 5, 054112, 2005.
135. J. A. Morrone and M. E. Tuckerman, *Ab initio molecular dynamics study of proton mobility in liquid methanol*, J. Chem. Phys., **117**, 4403, 2002.
136. Linas Vilciauskas, Mark E. Tuckerman, Gabriel Bester, Stephen J. Paddison, and Klaus-Dieter Kreuer, *The mechanism of proton conduction in phosphoric acid*, Nature Chemistry, **4**, no. 6, 461–466, June 2012.
137. Bernd Ensing, Alessandro Laio, Francesco L. Gervasio, Michele Parrinello, and Michael L. Klein, *A Minimum Free Energy Reaction Path for the E2 Reaction between Fluoro Ethane and a Fluoride Ion*, Journal of the American Chemical Society, **126**, no. 31, 9492–9493, Aug. 2004.
138. Jung Mee Park, Alessandro Laio, Marcella Iannuzzi, and Michele Parrinello, *Dissociation Mechanism of Acetic Acid in Water*, Journal of the American Chemical Society, **128**, no. 35, 11318–11319, Sept. 2006.
139. Xiandong Liu and Evert Jan Meijer, *Mechanism of Base-Promoted Dehydrochlorination of Pentachloroethane: Concerted or Stepwise?*, The Journal of Physical Chemistry A, **113**, no. 15, 3542–3544, Apr. 2009.
140. Xiandong Liu, Xiancai Lu, Evert Jan Meijer, Rucheng Wang, and Huiqun Zhou, *Acid dissociation mechanisms of Si(OH)4 and Al(H2O)63+ in aqueous solution*, Geochimica et Cosmochimica Acta, **74**, no. 2, 510–516, Jan. 2010.
141. Anna Pavlova and Evert Jan Meijer, *Understanding the Role of Water in Aqueous Ruthenium-Catalyzed Transfer Hydrogenation of Ketones.*, ChemPhysChem, **13**, no. 15, 3492–3496, Oct. 2012.
142. A. Curioni, M. Sprik, W. Andreoni, H. Schiffer, J. Hutter, and M. Parrinello, *Density functional theory-based molecular dynamics simulation of acid-catalyzed chemical reactions in liquid trioxane*, J. Am. Chem. Soc., **119**, 7218–7229, 1997.
143. E. J. Meijer and M. Sprik, *A Density Functional Study of the Addition of Water to SO3 in the Gas Phase and in Aqueous Solution*, J. Phys. Chem. A, **102**, 2893, 1998.
144. E. J. Meijer and M. Sprik, *Ab Initio Molecular Dynamics Study of the Addition Reaction of Water to Formaldehyde in Sulfuric Acid Solution*, J. Am. Soc. Chem., **120**, 6345, 1998.
145. M. Sprik, *Computation of the pK of liquid water using coordination constraints*, Chem. Phys., **258**, 139, 2000.
146. K. Doclo and U. Rothlisberger, *Conformational Equilibria of Peroxynitrous Acid in Water: a First-Principles Molecular Dynamics Study*, J. Phys. Chem. A, **104**, 6464–6469, 2000.
147. P.L. Geissler, C. Dellago, D. Chandler, J. Hutter, and M. Parrinello, *Autoionization in Liquid Water*, Science, **291**, 2121–2124, 2001.
148. B. Ensing, E. J. Meijer, P. E. Blochl, and E. J. Baerends, *Solvation Effects on the SN2 Reaction between CH3Cl and Cl- in Water*, J. Phys. Chem. A., **105**, no. 13, 3300–3310, 2001.
149. N. L. Doltsinis and M. Sprik, *Theoretical pKa estimates for solvated P(OH)5 from coordination constrained Car-Parrinello dynamics*, Phys. Chem. Chem. Phys., **5**, 2612, 2003.

150. F. Zipoli, M. Bernasconi, and A. Laio, *Ab initio simulations of Lewis-acid-catalyzed hydrosilylation of alkynes*, ChemPhysChem, **6**, 1772–1775, 2005.
151. A. Marco Saitta, Franz Saija, and Paolo V. Giaquinta, *Ab Initio Molecular Dynamics Study of Dissociation of Water under an Electric Field*, Physical Review Letters, **108**, no. 20, 207801, May 2012.
152. Antonino Marco Saitta and Franz Saija, *Miller experiments in atomistic computer simulations*, Proceedings of the National Academy of Sciences, p. 201402894, Sept. 2014.
153. Martin Konôpka, Robert Turanský, Matús Dubecký, Dominik Marx, and Ivan Stich, *Molecular Mechanochemistry Understood at the Nanoscale: Thiolate Interfaces and Junctions with Copper Surfaces and Clusters*, The Journal of Physical Chemistry C, **113**, no. 20, 8878–8887, May 2009.
154. Michiel Sprik and Giovanni Ciccotti, *Free energy from constrained molecular dynamics*, The Journal of Chemical Physics, **109**, no. 18, 7737–7744, Nov. 1998.
155. P. L. Geissler, C. Dellago, D. Chandler, J. Hutter, and M. Parrinello, *Ab initio analysis of proton transfer dynamics in (H2O)3H+*, Chem. Phys. Lett., **321**, 225–230, 2000.
156. Jerry B. Abrams, Mark E. Tuckerman, and Glenn J. Martyna, "Equilibrium statistical mechanics, non-hamiltonian molecular dynamics, and novel applications from resonance-free timesteps to adiabatic free energy dynamic", Chapter 5 of this book.
157. Michele Monteferrante, Sara Bonella, Simone Meloni, Eric Vanden-Eijnden, and Giovanni Ciccotti, *Calculations of free energy barriers for local mechanisms of hydrogen diffusion in alanates*, SCIENTIFIC MODELING AND SIMULATIONS, **15**, no. 1-3, 187–206, APR 2008.
158. Alessandro Laio and Michele Parrinello, *Escaping free-energy minima*, Proc. Nat. Acad. Sciences, **99**, 12562–12566, 2002.
159. Roman Martoňák, Alessandro Laio, Marco Bernasconi, Chiara Ceriani, Paolo Raiteri, Federico Zipoli, and Michele Parrinello, *Simulation of structural phase transitions by metadynamics*, Zeitschrift für Kristallographie, **220**, no. 5-6, 489–498, 2005.
160. Davide Branduardi, Francesco Luigi Gervasio, and Michele Parrinello, *From A to B in free energy space*, J. Chem. Phys., **126**, 054103, 2007.
161. Massimiliano Bonomi, Davide Branduardi, Giovanni Bussi, Carlo Camilloni, Davide Provasi, Paolo Raiteri, Davide Donadio, Fabrizio Marinelli, Fabio Pietrucci, Ricardo A Broglia, and Michele Parrinello, *PLUMED: A portable plugin for free-energy calculations with molecular dynamics.*, Computer Physics Communications, **180**, 1961–1972, 2009.
162. Fabio Pietrucci and Wanda Andreoni, *Graph Theory Meets Ab Initio Molecular Dynamics: Atomic Structures and Transformations at the Nanoscale*, Physical Review Letters, **107**, no. 8, 085504, Aug. 2011.
163. R. Vuilleumier and M. Sprik, *Electronic control using density functional perturbation methods*, Chem. Phys. Lett., **365**, 305, 2002.
164. Leonardo Guidoni and Ursula Rothlisberger, *Scanning Reactive Pathways with Orbital Biased Molecular Dynamics*, Journal of Chemical Theory and Computation, **1**, no. 4, 554–560, July 2005.
165. Jun Cheng, Marialore Sulpizi, and Michiel Sprik, *Redox potentials and pKa for benzoquinone from density functional theory based molecular dynamics*, The Journal of Chemical Physics, **131**, no. 15, 154504, Oct. 2009.
166. Jun Cheng, Xiandong Liu, Joost VandeVondele, Marialore Sulpizi, and Michiel Sprik, *Redox Potentials and Acidity Constants from Density Functional Theory Based*

Molecular Dynamics, Accounts of Chemical Research, **47**, no. 12, 3522–3529, Dec. 2014.

167. I. Tavernelli, R. Vuilleumier, and M. Sprik, *Ab initio molecular dymamics for molecules with variable numbers of electrons*, Phys. Rev. Lett., **88**, no. 21, 213002, 2002.
168. J. Blumberger, Y. Tateyama, and M. Sprik, *Ab initio molecular dynamics simulation of redox reactions in solution*, Comp. Pys. Comm., **169**, 256, 2005.
169. Sami Tazi, Benjamin Rotenberg, Mathieu Salanne, Michiel Sprik, and Marialore Sulpizi, *Absolute acidity of clay edge sites from ab-initio simulations*, Geochimica et Cosmochimica Acta, **94**, no. 0, 1–11, Oct. 2012.
170. Shawn M. Kathmann, I-Feng William Kuo, and Christopher J. Mundy, *Electronic Effects on the Surface Potential at the Vapor?Liquid Interface of Water*, Journal of the American Chemical Society, **130**, no. 49, 16556–16561, Dec. 2008.
171. Francesca Costanzo, Marialore Sulpizi, Raffaele Guido Della Valle, and Michiel Sprik, *The oxidation of tyrosine and tryptophan studied by a molecular dynamics normal hydrogen electrode*, The Journal of Chemical Physics, **134**, no. 24, 244508, June 2011.
172. Christopher Adriaanse, Jun Cheng, Vincent Chau, Marialore Sulpizi, Joost Vande-Vondele, and Michiel Sprik, *Aqueous Redox Chemistry and the Electronic Band Structure of Liquid Water*, The Journal of Physical Chemistry Letters, **3**, no. 23, 3411–3415, Dec. 2012.
173. M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, *Efficient and general algorithms for path integral Car-Parrinello molecular dynamics*, J. Chem. Phys., **104**, 5579–5588, 1996.
174. D. Marx, M. E. Tuckerman, and G. J. Martyna, *Quantum dynamics via adiabatic ab initio centroid molecular dynamics*, Comput. Phys. Commun., **118**, 166–184, 1999.
175. Mark E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, Oxford Graduate texts. Oxford University Press, 2010.
176. M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, *On the Quantum Nature of the Shared Proton in Hydrogen Bonds*, Science, **275**, 817–820, 1997.
177. S. Miura, M.E. Tuckerman, and M.L. Klein, *An ab initio path integral molecular dynamics study of double proton transfer in the formic acid dimer*, J. Chem. Phys., **109**, 5290, 1998.
178. D. Marx, M. E. Tuckerman, and M. Parrinello, *Solvated excess protons in water: quantum effects on the hydration structure*, J. Phys.: Condens. Matter A, **12**, 153–159, 2000.
179. Michele Ceriotti, Jérôme Cuny, Michele Parrinello, and David E. Manolopoulos, *Nuclear quantum effects and hydrogen bond fluctuations in water*, Proceedings of the National Academy of Sciences, **110**, no. 39, 15591–15596, Sept. 2013.
180. Joseph A. Morrone, Varadharajan Srinivasan, Daniel Sebastiani, and Roberto Car, *Proton momentum distribution in water: an open path integral molecular dynamics study*, The Journal of Chemical Physics, **126**, no. 23, 234504, June 2007.
181. Michele Ceriotti and David E. Manolopoulos, *Efficient First-Principles Calculation of the Quantum Kinetic Energy and Momentum Distribution of Nuclei*, Physical Review Letters, **109**, no. 10, 100604, Sept. 2012.
182. Giovanni Romanelli, Michele Ceriotti, David E. Manolopoulos, Claudia Pantalei, Roberto Senesi, and Carla Andreani, *Direct Measurement of Competing Quantum*

Effects on the Kinetic Energy of Heavy Water upon Melting, The Journal of Physical Chemistry Letters, **4**, no. 19, 3251–3256, Oct. 2013.
183. Lin Lin, Joseph A. Morrone, Roberto Car, and Michele Parrinello, *Momentum distribution, vibrational dynamics, and the potential of mean force in ice*, Physical Review B, **83**, no. 22, 220302, June 2011.
184. N. L. Doltsinis and D. Marx, *Nonadiabatic Car-Parrinello molecular dynamics*, Phys. Rev. Lett., **88**, 166402, 2002.
185. H. Langer and N. L. Doltsinis, *Excited state tautomerism of the DNA base guanine: A restricted open-shell Kohn-Sham study*, J. Chem. Phys., **118**, 5400–5407, 2003.
186. N. L. Doltsinis, *Excited state proton transfer and internal conversion in o-hydroxybenzaldehyde: new insights from nonadiabatic ab initio molecular dynamics*, Mol. Phys., **102**, 499–506, 2004.
187. N. L. Doltsinis and D. S. Kosov, *Plane wave/pseudopotential implementation of excited state gradients in density functional linear response theory: A new route via implicit dierentiation*, J. Chem. Phys., **122**, 144101, 2005.
188. Enrico Tapavicza, Ivano Tavernelli, and Ursula Rothlisberger, *Trajectory Surface Hopping within Linear Response Time-Dependent Density-Functional Theory*, Physical Review Letters, **98**, no. 2, 023001, Jan. 2007.
189. Ivano Tavernelli, Basile F. E. Curchod, Andrey Laktionov, and Ursula Rothlisberger, *Nonadiabatic coupling vectors for excited states within time-dependent density functional theory in the Tamm–Dancoff approximation and beyond*, The Journal of Chemical Physics, **133**, no. 19, 194104, Nov. 2010.
190. Basile F. E. Curchod, Ursula Rothlisberger, and Ivano Tavernelli, *Trajectory-Based Nonadiabatic Dynamics with Time-Dependent Density Functional Theory*, ChemPhysChem, **14**, no. 7, 1314–1340, May 2013.
191. P. López-Tarifa, M.-A. Hervé du Penhoat, R. Vuilleumier, M.-P. Gaigeot, I. Tavernelli, A. Le Padellec, J.-P. Champeaux, M. Alcamí, P. Moretto-Capelle, F. Martín, and M.-F. Politis, *Ultrafast Nonadiabatic Fragmentation Dynamics of Doubly Charged Uracil in a Gas Phase*, Physical Review Letters, **107**, no. 2, 023202, July 2011.
192. Pablo López-Tarifa, Marie-Pierre Gaigeot, Rodolphe Vuilleumier, Ivano Tavernelli, Manuel Alcamí, Fernando Martín, Marie-Anne Hervé du Penhoat, and Marie-Françoise Politis, *Ultrafast Damage Following Radiation-Induced Oxidation of Uracil in Aqueous Solution*, Angewandte Chemie International Edition, **52**, no. 11, 3160–3163, 2013.

High-Dimensional Neural Network Potentials as a Tool to Study Solvation[1]

Jörg Behler

Lehrstuhl für Theoretische Chemie,
Ruhr-Universität Bochum, D-44780 Bochum, Germany

E-mail: joerg.behler@theochem.ruhr-uni-bochum.de

A lot of progress has been made in recent years in the development of atomistic potentials employing machine learning (ML) techniques. In contrast to most conventional potentials, which are based on physical approximations and simplifications to derive an analytic functional relation between the atomic configuration and the potential-energy, ML potentials rely on simple but very flexible mathematical terms without a direct physical meaning. Instead, in case of ML potentials the topology of the potential-energy surface is "learned" by adjusting a number of parameters with the aim to reproduce a set of reference electronic structure data as accurately as possible. Due to this bias-free construction they are applicable to a wide range of systems without changes in their functional form, and a very high accuracy close to the underlying first-principles data can be obtained. Neural network potentials (NNPs), which have first been proposed about two decades ago, are an important class of ML potentials. While the first NNPs have been restricted to small molecules with only a few degrees of freedom, they are now applicable to high-dimensional systems containing thousands of atoms, which enables addressing a variety of problems in chemistry, physics and materials science. In this lecture the basic ideas of NNPs are presented with a special focus on constructing NNPs for high-dimensional condensed systems. A recipe for the construction of these potentials is given and remaining limitations of the method are discussed.

1 Introduction

An accurate description of the atomic interactions is of vital importance for carrying out reliable computer simulations in chemistry, physics and materials science. Employing the Born-Oppenheimer approximation[2], various electronic structure methods are available to directly calculate the potential-energy and the atomic forces for a given nuclear configuration based on the laws of quantum mechanics. Often, the potential-energy of interest is the ground state energy, although the methods discussed below can equally be applied to excited state energies if they are available. The optimum choice of the electronic structure method depends on the system and typically requires to find an acceptable compromise between efficiency and accuracy for the problem of interest, as the exact solution of the Schrödinger equation is impossible for essentially all relevant problems. Each of these electronic structure calculations then provides a specific point on the multidimensional potential-energy surface (PES), which is in general a real-valued function depending on all atomic coordinates in the system yieling its potential-energy. In this lecture, the terms PES and "potential" will be used synonymously for this function.

For practical reasons the number of electronic structure energies that can be calculated and stored is limited. Therefore, even in large data bases the energies and forces of most configurations that are visited, e.g., in molecular dynamics (MD) simulations will not be available beforehand. Consequently, in *ab initio* MD[3,4] the energies and forces need to

be calculated "on-the-fly", typically employing density-functional theory (DFT)[5,6]. Alternatively, an analytic expression for the PES can be constructed and used in the simulations, which allows to perform MD simulations more efficiently as the evaluation of such expressions is much faster than solving the quantum mechanical equations. Also this approach is justified by the Born-Oppenheimer approximation, because the Hamiltonian and therefore also the ground state potential-energy is completely defined by the atomic positions, the nuclear charges, and the total charge of the system. Consequently, in principle a well-defined relation between the atomic structure and its potential-energy exists. Unfortunately, in most cases the corresponding functional form is too complicated to be derived analytically.

A pragmatic solution to this problem it the introduction of approximate PESs, and there are two fundamental approaches. In the conventional approach the solution of the Schrödinger equation is replaced by using a simplified energy expression based on physical considerations and reasonable approximations. The accuracy of these "physical potentials", which represent the vast majority of potentials in the literature, is thus limited by the imposed fixed functional form. Still, in most cases the essential features of the PES are described correctly, and reliable simulations can be carried out. There are many different physical potentials of varying form and complexity, ranging from classical force fields[7–11] in biochemistry to reactive potentials in materials science[12–14].

The central idea of classical force fields is the decomposition of the total energy into low-dimensional bonding two-, three- and four-body terms representing covalent bonds, bonding angles and dihedral angles. In addition, electrostatic and van der Waals interactions given by Coulombs law and the Lennard-Jones potential, respectively, are used. In the low-dimensional bonding terms, only the immediate environment is taken into account. Since this approach would not be feasible based on the specification of the chemical elements alone, as the bonding properties are strongly influenced by the positions of the neighboring atoms, additional information on the bonding properties of the atoms is included via the introduction of atom types classifying the atoms according to functional groups. In this way, total energy expressions with low-dimensional, approximately additive energy terms can be obtained. The most severe limitation of classical force fields, i.e., their inability to describe the making and breaking of bonds, has been overcome in several reactive force fields, e.g. in the ReaxFF method[15].

Most atomistic potentials in the field of materials science are based on the concept of the bond order[16] to take into account the effect of the atomic environments on the bonding properties. This is of fundamental importance, as in many problems of materials science involving systems like metals and alloys a good description of very different atomic environments is mandatory due to the importance of many-body contributions to the potential-energy, which render a decomposition of the potential into individual low-dimensional terms impossible. Further, simulations of these phenomena involve significant atomic rearrangements and thus require reactive potentials without the definition of fixed atom types.

An alternative class of atomistic potentials employs very flexible functions without a direct physical meaning. The aim of these "mathematical potentials" or machine learning potentials is to fit an analytic expression to a set of reference data obtained in electronic structure calculations as accurately as possible. These potentials are not yet widely distributed, but a number of promising approaches have been proposed in recent years differing in the types of functions that are used. They comprise permutation invariant

polynomials[17,18], the modified Shepard method employing Taylor expansions[19,20], Gaussian processes[21–25], interpolating moving least squares[26,27], artificial neural networks[28,29] and also support vector machines[30]. A more detailed discussion and comparison of these approaches can be found elsewhere[31]. While the construction and use of these potentials has to be done with care to ensure that the correct physical shape of the PES is obtained, they are numerically extremely accurate and can be constructed even in difficult cases, for which no reasonable approximate physical potential can be found.

Physical as well as mathematical potentials provide the energy directly as a function of the atomic positions using well-defined analytic functions. Therefore, they both represent "atomistic potentials", and the use of such potentials enables the calculation of energies and forces many orders of magnitude faster than electronic structure methods. They are particularly useful

- if long MD trajectories or extended Monte Carlo simulations are required
- if many MD trajectories are needed to obtain statistically converged results

and/or

- if the systems are too large for the application of electronic structure methods

In all these cases, they allow to extend the time and length scales of computer simulations beyond the realm of *ab initio* methods provided that a suitable functional form and a reliable set of parameters can be found, which is often a substantial challenge.

Mathematical potentials like the neural network potentials (NNPs) discussed in this lecture are capable of describing even very complex PESs, for which physical potentials are difficult or even impossible to derive. Typical candidates for applications of such potentials are systems containing

- many different types of interactions, like covalent, ionic and metallic bonding as well as dispersion interactions
- "unusual" atomic environments, e.g. in amorphous systems, structures emerging during phase transitions or being present in coordination chemistry
- complicated reaction and transition paths involving the making and breaking of bonds

Artificial neural networks (NNs)[32,33] have been introduced in 1943 to model and understand the signal processing in the brain[34], and in the following decades they have found wide use in many fields of science[35] due to their pattern recognition and data classification capabilities. There are many types of NNs with different functional forms[32,33]. A general definition covering all these types has been given by Kohonen[36]:
"*Artificial neural networks are massively parallel interconnected networks of simple (usually adaptive) elements and their hierarchical organizations which are intended to interact with the objects of the real world in the same way as biological nervous systems do.*"

In particular multilayer feed-forward (MLFF) NNs have been demonstrated to be useful for the construction of PESs due to their ability to represent arbitrary functions. It has been proven independently by a number of researchers that MLFF NNs are "universal approximators"[37–41], i.e., they enable approximating unknown multidimensional functions to an, in principle, arbitrary accuracy based on a set of known function values:

"*[...] standard multilayer feedforward networks with as few as one hidden layer using arbitrary squashing functions are capable of approximating any [...] function from one finite dimensional space to another to any desired degree of accuracy, provided sufficiently many hidden units are available. In this sense, multilayer feedforward networks are a class of universal approximators.*"[38]

This is the theoretical foundation for employing NNs to construct atomistic potentials. However, in practice this formal proof is of limited use as neither the number or distribution of training points nor the required size of the NN for the representation of these points is known. Still, the remarkable result is that there is no fundamental restriction in the accuracy that can be achieved when constructing NNPs, which is an important difference to physical potentials, which have intrinsic limitations due to their rather inflexible functional forms.

In general, NNPs can be defined by three criteria:[42]

1. NNPs provide a direct functional relation between the atomic configuration and the potential-energy employing one or more artificial NNs.

2. NNPs are constructed using a set of first-principles data, usually total energies and sometimes energy derivatives, obtained from a single electronic structure method. Experimental data cannot be used as experiments often have uncertainties being substantially larger than the fitting errors of NNPs. Results from different electronic structure methods cannot be combined as numerically inconsistent information must be avoided.

3. NNPs do not contain any approximations apart from the intrinsic limitations of the chosen reference electronic structure method. No other *ad hoc* empirical functional components are included.

NNPs offer a number of advantages for the construction of PESs:

- Energies can be fitted to high accuracy with very small remaining errors compared to the underlying reference data.

- NNPs can be calculated efficiently and require much less CPU time than electronic structure calculations.

- No knowledge about the functional form of the PES is required.

- The NN energy expression is unbiased, generally applicable to all types of bonding and does not require system-specific modifications.

Still, there are also disadvantages of NNPs that one should be aware of:

- The evaluation of NNPs is notably slower than the use of simple classical force fields.

- NNPs have no physical basis and only very limited extrapolation capabilities. Therefore they can fail spectacularly if they are not used properly.

- The construction of NNPs requires substantial effort, and a large number of training points from demanding electronic structure calculations is required.

- Currently, NNPs are limited either to systems containing only a few elements or a small number of atoms.

In the past two decades, NNPs have been constructed for many types of systems. Initially, they have been restricted to small molecules[43–47] and systems, whose complexity has been strongly reduced by freezing a majority of the degrees of freedom, like small molecules interacting with frozen metal surfaces[28,48,49,29,50–59], or extensions of water potentials to include polarzation effects truncated at low order while keeping rigid water monomer geometries[60,61]. A comprehensive summary of previous work on NNPs can be found in two recent reviews[62,63].

Only a few attempts have been made to date to construct NNPs for high-dimensional systems. The term "high-dimensional" has been used frequently and with different meanings in the literature. Here, we consider a potential as being high-dimensional if it is applicable to systems containing thousands of atoms and all their degrees of freedom explicitly. One early approach to construct high-dimensional NNPs by Smith and coworkers published in 1999[64,65] is based on a decomposition of the system into chains of atoms of increasing length, whose energies are constructed using a NN of variable size. Surprisingly, this promising idea has not been further developed until 2007, when it was extended to represent the PES of silicon using tight binding reference data[66]. Since then, the only application of this method has been the construction of the DFT PES of the same system in 2008[67]. Another very appealing approach introduced by Manzhos and Carrington is based on a rigorous many-body expansion of the potential-energy using redundant coordinates and a high-dimensional model representation[68–72]. Here, in essence, the potential-energy is expressed as a sum of terms of increasing order in the spirit of a many-body expansion, and it has been demonstrated that formally low-dimensional NNs can be used if the input coordinates are constructed as a function of the original atomic coordinates. This method, which is very accurate and has no fundamental conceptual limitations, has been applied to date only to rather small organic molecules, as the number of NN evaluations increases rapidly with system size making it very costly for larger systems. Still, it is probably the most systematic approach to construct NNPs.

In this lecture the basic methodology of a NNP approach, which has been introduced by Behler and Parrinello[73,74,42] to address high-dimensional systems, and its extensions will be discussed with a focus on its construction and applicability. To date NNPs based on this approach have been constructed for various systems including silicon[73,75,76], copper[77], carbon[78,79], sodium[80,81], zinc oxide[82], copper clusters supported at zinc oxide[83], neutral[84–86] as well as protonated water clusters[87], and the phase change material GeTe[88–91]. Many of the techniques described below are very general and can be applied to all types of NNPs and even beyond to other types of ML potentials.

2 Conventional Neural Network Potentials

Most conventional NNPs use a single MLFF NN to construct a direct functional relation between the atomic configuration and the potential-energy. For this purpose, a number of artificial neurons is organized in several layers as shown in Fig. 1. The potential energy E is obtained in the neuron, or node, in the output layer. It depends on the atomic configuration that is provided to the NN in form of a vector of input coordinates $\mathbf{G} = \{G_i\}$. The specific

choice of input coordinates is a crucial aspect of any NNP, as these coordinates need to fulfill a number of requirements. In between the input and the output layer there are one or more so-called "hidden layers". As their name implies, they have no physical meaning but define the functional form of the NN and provide the required mathematical flexibility to construct the functional relation between the input and the output of the NN. The more hidden layers are used and the more nodes are included in each hidden layer, the more flexible is the functional form of the NN. A typical NN architecture contains two to three hidden layers and up to typically about 50 nodes per layer. The entity of all layers including the input and output layer as well as the number of nodes per layer defines the architecture of the NN. For instance, the NN shown in Fig. 1 can be described by the short notation 3-5-5-1, where each number specifies the number of nodes in one layer starting from the input layer.

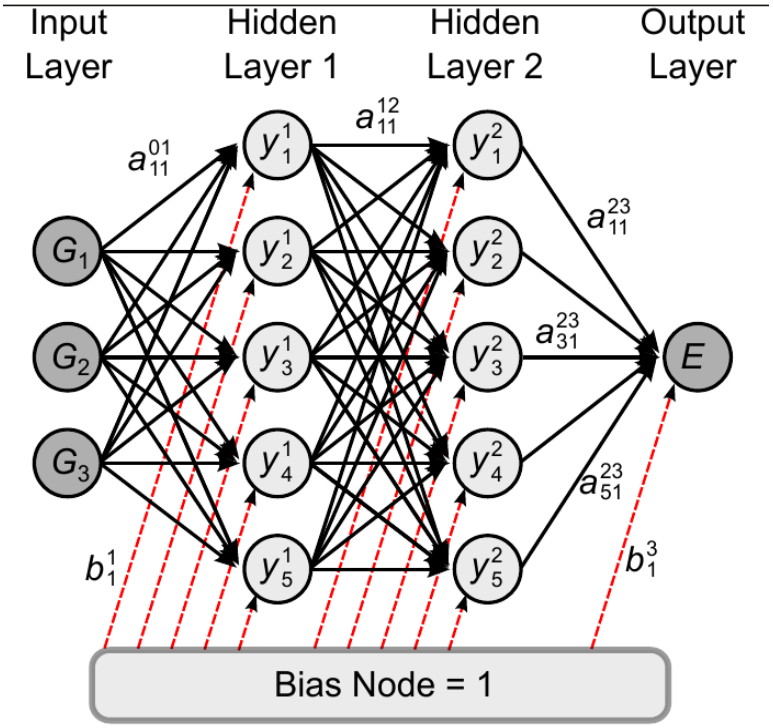


Figure 1: Schematic structure of a small feed-forward neural network (NN). The nodes are arranged in layers and the goal is to establish a functional relation between the energy output E and the structure described by a vector of input coordinates $\{G_i\}$. The functional form of the NN is given by Eq. 2.

As indicated by the arrows in Fig. 1 each node in each layer is connected to the nodes in the adjacent layers by a weight parameters, where the symbol a_{ij}^{kl} is used for the parameter connecting node i in layer k with node j in layer l with $l = k + 1$. The input layer has the superscript 0 and the arrows label the direction of the flow of information through the

network. Further, there is a bias node connected to all nodes in the hidden layers and to the output node providing the number "one", which can be scaled by a bias weight b_i^j, where j is the layer of the target node and i is its number. The value y_i^j of any node i in any hidden layer j can then be calculated from the values of the N_j nodes in the previous layer and the connecting and bias weight parameters by

$$y_i^j = f_i^j \left(b_i^j + \sum_{k=1}^{N_{j-1}} a_{k,i}^{j-1,j} \cdot y_k^{j-1} \right) \quad . \tag{1}$$

This equation is essentially a linear combination of the values of the nodes in the previous layer, which has been shifted by the bias weight. As the topology of a PES generally cannot be expressed as a simple linear combination, a non-linear function f_i^j is applied, which is called "activation function" or "basis function" of the NN. It has the purpose to provide the capability of fitting arbitrary functions. As part of the NN total energy expression the activation functions must be differentiable to determine analytic forces and to calculate the derivatives of the NN output with respect to the weight parameters needed for the gradient-based optimization of the weights. In case that forces are used for the NN training as well, also second derivatives of the activation functions are needed.

There are many possible choices for the activation functions. They all have in common that they have a non-linear region and that they typically converge to 1, -1 or 0 for very large and very small arguments. Frequent choices are the hyperbolic tangent, the sigmoid function and Gaussians. Still, also periodic activation functions like trigonometric functions[92] and exponentials[69] have been proposed. For the node in the output layer, often a linear function is used to avoid any restriction in the range of possible output values of the NN. The complete functional form of the small example NN in Fig. 1 is a hierarchy of nested activation functions acting on linear combinations of the values in the previous layer given by

$$E = f_1^3 \left(b_1^3 + \sum_{k=1}^{5} a_{k1}^{23} \cdot f_k^2 \left(b_k^2 + \sum_{j=1}^{5} a_{jk}^{12} \cdot f_j^1 \left(b_j^1 + \sum_{i=1}^{3} a_{ij}^{01} \cdot G_i \right) \right) \right) \quad . \tag{2}$$

Starting from the left side of Fig. 1 first the values of the nodes in the first hidden layer are calculated, followed by the nodes in the second hidden layer and so forth until the output node is reached. The NN energy expression also provides access to forces and higher derivatives. These should not be fitted separately using additional output nodes as this necessarily would result in numerical inconsistencies between the energies and forces. Instead the forces can be calculated as analytic derivatives as discussed below.

In order to obtain physically meaningful energy outputs, NNPs need to be trained to a set of known reference points obtained from electronic structure calculations. This is done by optimizing the weight parameters, which are initially chosen as random numbers, iteratively to minimize the error of this training set. This is often a very time-consuming step, as the training data sets can become very large and also the number of weight parameters N_w to be optimized can be substantial,

$$N_\mathrm{w} = \sum_{k=1}^{M_\mathrm{HL}+1} \left(N_{k-1} \cdot N_k + N_k \right), \tag{3}$$

where M_{HL} is the number of hidden layers and N_k is the number of nodes in layer k including the output layer. NNPs containing only a single feed-forward NN have been constructed successfully for a number of low-dimensional systems containing up to about 12 degrees of freedom. A highly recommended tutorial-style discussion of fitting simple model functions by NNs can be found elsewhere[93].

3 High-Dimensional Neural Network Potentials

Using a single feed-forward NN for the construction of high-dimensional NNPs is impossible for several reasons. First of all, for each atom added to the system three new degrees of freedom need to be included in form of additional input nodes. If the total number of input nodes becomes too large, the construction of the NNP will be inefficient because there are too many weight parameters to be determined in the fitting process. Further, in applications of the NNP the calculation of the output energy will be slow. Finally, constructing a sufficiently dense set of training points becomes very challenging in a too high-dimensional coordinate space both in terms of selecting the points and in terms of CPU time.

Another fundamental problem, which is present even for small molecules, is related to the symmetry of the NN. For a water monomer, which can be described by three interatomic distances, both OH bonds are chemically equivalent. If the positions of both hydrogen atoms are exchanged, the atomic configuration must still have the same potential-energy since chemically the structure has not changed. As the connecting weights of a NNP all have numerically different values, changing the order of the input coordinates, i.e. of the two OH-bond lenghts, will result in a numerically different energy output. This problem has been recognized very early, and a solution has been proposed already in 1998 by Gassner et al.[61]. Instead of using a set of interatomic distances a symmetrization step corresponding to a transformation of the initial internal coordinates is introduced in this approach. Consequently, chemically equivalent atoms lose their identity and all possible representations, i.e. orders of the atoms, yield the same NN input vector. Therefore, the same output energy is obtained as it should be. Unfortunately, this method, which has also been extended to the scattering of molecules at solid surfaces[49], is applicable only to very small systems, as the complexity of the symmetrization step increases rapidly with the number of atoms. Even this simple example illustrates the crucial importance of the choice of input coordinates, and this topic will be discussed in more detail in below.

Finally, any NNP consisting of a single NN is only applicable to the system size that has been used for its construction, which is another severe problem. If an atom would be added when applying the potential, the connecting weights for the additional input nodes would not be available. If, on the other hand, an atom would be removed, its NN input coordinates would be ill-defined, and using simply a set of zero input values for these nodes is not possible because these numbers would change in frequently used preprocessing steps of the input coordinates (cf. below).

A solution to all these problems is to construct the energy E_{s} of the system as a sum of N_{atom} atomic energy contributions E_i, which are provided by a set of individual atomic NNs,

$$E_{\mathrm{s}} = \sum_{i=1}^{N_{\mathrm{atom}}} E_i. \tag{4}$$

The atomic energy contributions depend on the local chemical environments up to a cutoff

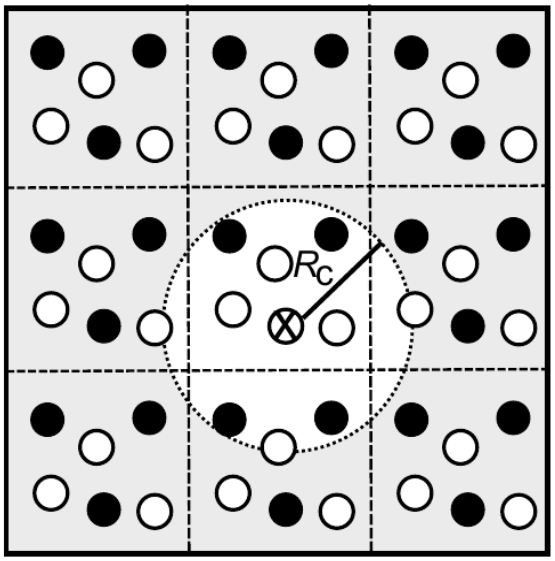


Figure 2: Illustration of the atomic environment defined by a cutoff radius R_c in a periodic system containing two different elements (white and black circles). In high-dimensional neural network potentials, all atoms inside the cutoff sphere determine the energy contribution of the central atom to the potential-energy.

radius R_c as shown in Fig. 2. This cutoff radius, which typically has values between 6 and 10 Å, is a convergence parameter and needs to be tested to ensure that all energetically relevant interactions are included. The positions of the neighboring atoms within the resulting cutoff sphere are then described by a set of many-body symmetry functions. They serve as a structural fingerprint and are supplied to the atomic NNs as a vector containing usually between 20 and 100 function values. The structure of the atomic NNs is essentially identical to the structure of the single-NN NNPs discussed in the previous Section (cf. Fig. 1) with the only difference that the output is now just a part of the total energy.

The resulting high-dimensional NN approach is shown schematically in Fig. 3 for a system consisting of N water molecules. For each atom there is a separate line starting from the Cartesian coordinate vector $\mathbf{R}$ of the atom. In the next step a vector of many-body symmetry functions $\mathbf{G}$ is constructed for each atom, which describes the arrangement of all atoms in the chemical environment. Consequently, the symmetry function values depend on the Cartesian coordinates of all other atoms as indicated by the grey arrows. This vector of symmetry function values is the input for the atomic NN yielding the energy contribution of the atom. Once each atomic energy contribution has been determined, the total energy E_s is obtained by summing all contributions.

It should be noted that although the atomic NNs provide "atomic energies", no information about partitioned atomic energies is required and the weight parameters can be determined using total energies from electronic structure calculations only. The condition of permutation invariance of the total energy with respect to the order of the atoms is

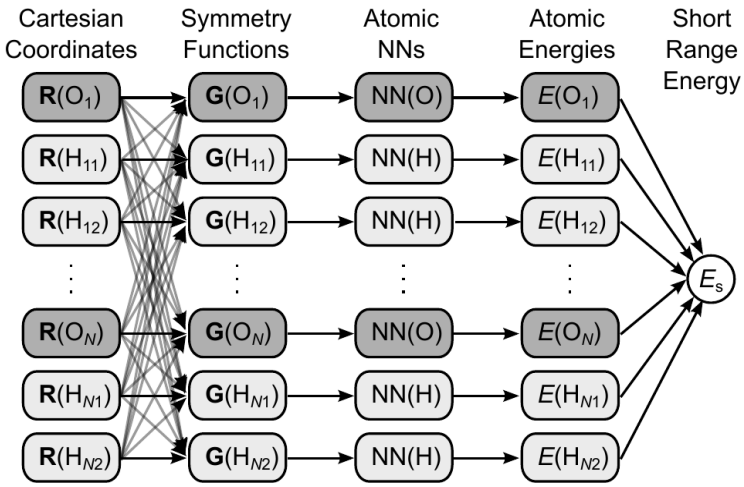


Figure 3: Schematic structure of a high-dimensional neural network potential for a system containing N water molecules. First, for each atom the Cartesian coordinate vectors $\mathbf{R}$ are transformed to symmetry function vectors $\mathbf{G}$, which describe the atomic environments and serve as input for the atomic neural networks providing the atomic energies. Finally, the atomic energy contributions are added to yield the short range energ E_s.

automatically satisfied in this method as the order of the summation does not change the resulting total energy. For each element, like oxygen or hydrogen, there is a separate type of atomic NN with its own architecture and set of weight parameters, but for a given element all atomic NNs are identical. Different atomic energy contributions arise only because of the different NN input vectors reflecting the different environments. As a result, the high-dimensional NN method is also applicable to systems of different size, in the training of the NNP as well as in its application. If an atom is added, the scheme in Fig. 3 is extended by another line of the respective element. If an atom is removed from the system, its line is deleted.

In summary, by using the energy expression in Eq. 4 the most severe problems of conventional NNPs can be solved:

- The introduction of a cutoff reduces the effective dimensionality to the energetically relevant interactions, which allows to use NNs of tractable size. Still, the remaining atomic environments often contain up to 100 neighboring atoms or more enabling the inclusion of high-order many-body effects.

- The total energy is invariant with respect to the order of the atoms.

- The NNP is applicable to systems containing different numbers of atoms without the need to determine a different set of weight parameters for each system size.

In the original high-dimensional NNP method of Behler and Parrinello[73] all interactions between atoms separated by more than the cutoff radius are truncated. Still, there might be interactions beyond this radius that should be included, like rather long-ranged electrostatic interactions. For this purpose, it is possible to extend the total energy expres-

sion in Eq. 4 by an additional electrostatic term[82,84] E_{elec},

$$E = E_s + E_{\mathrm{elec}} = \sum_{i=1}^{N_{\mathrm{atom}}} E_i + E_{\mathrm{elec}}. \tag{5}$$

The electrostatic energy is calculated using either Coulombs law or an Ewald summation[94], i.e. no spatial truncation is applied. The required charges are constructed as environment-dependent atom-centered charges expressed by a second set of atomic NNs in the same way as the (now "short-ranged") energy contributions in Fig. 3. Like atomic energies, atomic charges are no quantum mechanical observables, but still applying charge partitioning methods is much more common than using energy partitioning schemes[95,96]. We have therefore chosen to use Hirshfeld charges[97] obtained in electronic structure calculations to determine the weight parameters of the atomic charge NNs. In principle, however, it would also be possible to couple the fitting of the weight parameters of the short-range energy and charge NNs and to determine the charges under the only constraint that the best possible total energy is obtained. This, however, is computationally substantially more demanding. Also higher electrostatic multipoles can be represented by NNs, as has been demonstrated by Popelier and coworkers[98,99].

It should be noted, however, that according to our experience including long-range electrostatic interactions does only marginally improve the description of the PES for most systems as all electrostatic interactions within the cutoff region can be described as well by the short range part of the NNP, which does not distinguish between electrostatic and non-electrostatic interactions. Consequently, the electrostatic NN is only relevant for electrostatic interactions of atoms separated by more than the cutoff radius. Including electrostatics explicitly requires to separate the reference total energies into an electrostatic and a short range contribution before the fitting to avoid a double counting of the electrostatic part. Because of the singularity of the Coulomb interaction at short distances, removing the electrostatic energy from the total energy can thus even give rise to an increased corrugation of the remaining short range part of the potential. This complicates the fitting and can be avoided by screening the electrostatic interactions at short distances[84]. This screening has no consequences for the accuracy of the potential as the screened part of the electrostatic energy is by definition compensated by the short-range energies.

Finally, also van der Waals terms can be added to NNPs. Conceptually, there is no need to add these terms as all kinds of interactions are treated by NNPs on an equal footing and the most relevant part of van der Waals interactions is acting inside the cutoff radius. Still, DFT calculations are frequently used to train NNs and depending on the choice of the exchange correlation functional DFT often provides a poor description of van der Waals interactions, which is a well-known deficiency of GGA functionals that has received a lot of attention in recent years[100]. A number of correction schemes have been proposed to improve the description of van der Waals interactions. In case of NNPs in particular the D3 method by Grimme[101] has been found to be useful[85], because the NN energies are numerically very close to the DFT energies. Consequently, the D3 correction can equally be applied to the NN energies after the evaluation of the NNP. Alternatively, the corrections can be applied to the training set before fitting the NN parameters, which reduces the computational costs of the NNP to that of standard high-dimensional NNPs and enables using other correction schemes[102].

4 Symmetry Functions

4.1 Properties of Symmetry Functions

Having established the structure of high-dimensional NNPs, one of the most important decisions to be made concerns the description of the atomic configurations using a suitable set of coordinates. Cartesian coordinates cannot be used at all, because their numerical values are not invariant with respect to translations and rotations of the system. Since, on the other hand, the structure of a system and thus its energy remain unchanged upon these operations, so must be the set of input coordinates. In case of NNPs, many different types of coordinates have been used over the years like interatomic distances[60], symmetrized distances[61], functions including the symmetry of surfaces[29,49], complex combinations of coordinates in a many-body expansion[70], and many others.

In this lecture the transformed coordinates, which are many-body functions of all atomic positions inside the cutoff spheres, will be called "symmetry functions" for historic reasons. The term "symmetry" does not refer to the point group or the space group of the system, but to the basic requirement that structurally equivalent representations of the system must give rise to the same set of coordinate values. The most common but also the most critical operation is the permutation of atoms, i.e. their order in the structural description, since the order of the atoms is arbitrary. Any exchange of atoms of the same element, which is not at all restricted to positions being equivalent by symmetry, must not affect the vector of symmetry function values. Constructing such permutation invariant input coordinates, which is a common problem for all ML potentials[103], has been a major challenge in the advent of NNPs. Symmetries in the traditional meaning of symmetry operations do not play an important role, since they are included as a special case in the permutation symmetry and since in most applications like MD simulations at finite temperatures systems possess only C_1 symmetry.

For high-dimensional NNPs employing a cutoff radius to define the energetically relevant atomic environment it must be ensured that the number of symmetry functions does not depend on the number of atoms inside the cutoff sphere, as this can change in the course of MD simulations. Another obvious criterion for the choice of the symmetry functions is their ability to distinguish different atomic environments and to provide similar coordinate values for similar atomic configurations to facilitate the fitting process. Contradictory data, which emerge if two different structures with different energie have closely resembling symmetry function values, must be avoided as this strongly hampers the determination of the NN weight parameters. Finally, we note that even complicated functional forms of the symmetry functions are acceptable, since the transformation of the coordinates does not need to be inverted. In both cases, in the training of the NNP as well as in its application, the transformation is starting from the Cartesian coordinates to obtain the symmetry functions. A transformaiton in the opposite direction is not required.

In summary, a set of symmetry functions for the construction of high-dimensional NNPs must have the following properties:

- rotational and translational invariance
- invariance with respect to the permutation of atoms of the same element
- provide a unique description of the atomic positions

- constant number of function values independent of the number of atoms in the cutoff spheres

In the following Section, the functional forms of several types of symmetry functions suitable for the construction of high-dimensional NNPs will be discussed. Further details and additional functions can be found elsewhere[74].

4.2 The Functional Form of the Symmetry Functions

An important component of all symmetry functions is the cutoff function f_c defining the atomic environments. Two possible choices are the function

$$f_{\mathrm{c},1}\left(R_{ij}\right) = \begin{cases} 0.5 \cdot \left[\cos\left(\frac{\pi R_{ij}}{R_\mathrm{c}}\right) + 1\right] \text{ for } R_{ij} \leq R_\mathrm{c} \\ 0.0 \text{ for } R_{ij} > R_\mathrm{c}, \end{cases} \tag{6}$$

which is essentially the monotonically decreasing part of a cosine function whose domain has been rescaled from π to R_c and the function

$$f_{\mathrm{c},2}\left(R_{ij}\right) = \begin{cases} \tanh^3\left[1 - \frac{R_{ij}}{R_\mathrm{c}}\right] \text{ for } R_{ij} \leq R_\mathrm{c} \\ 0.0 \text{ for } R_{ij} > R_\mathrm{c}. \end{cases} \tag{7}$$

Both functions decrease with increasing distance R_{ij} between the central atom i and its neighbor j, which reflects qualitatively the reducing strength of the interactions between the atoms. At the cutoff radius R_c both functions have zero value and slope as shown in Fig. 4. This is important for the calculation of the forces, which require the determination of the derivatives of the symmetry functions with respect to the atomic Cartesian coordinates. Additionally, in case of cutoff function $f_{\mathrm{c},2}\left(R_{ij}\right)$ also the second derivative is zero at R_c. Formally, this is important to avoid discontinuities in the change of the forces when atoms enter or leave the cutoff spheres in MD simulations as this can possibly result in problems to achieve energy conservation. We found that as long as sufficiently large cutoff radii are used, i.e. 6 Å or more, also function $f_{\mathrm{c},1}\left(R_{ij}\right)$ provides an excellent energy conservation. In the symmetry functions discussed in the remaining part of this Section, the function f_c can be either of these two functions.

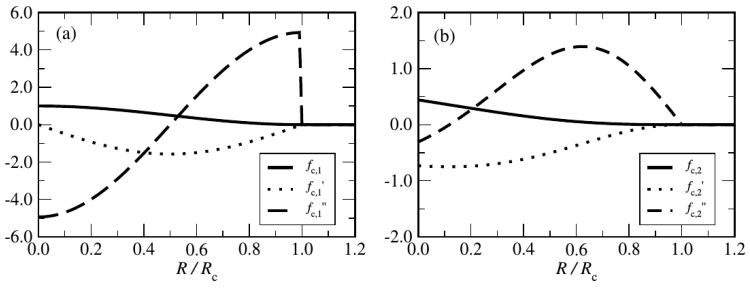


Figure 4: Plot of the cutoff functions $f_{\mathrm{c},1}$ and $f_{\mathrm{c},2}$ and their derivatives as defined in Eqns. 6 and 7, respectively.

Based on the cutoff function as a central component, several types of many-body symmetry functions can be constructed. There are two classes of symmetry functions: "radial"

symmetry functions, which describe the radial distribution of neighbors up to the cutoff radius, and "angular" symmetry functions specifying their angular arrangement. All symmetry functions have in common that each of them depends on the positions of all atoms inside the cutoff spheres. Therefore, in contrast to internal coordinates like bond lengths, their numerical values are not always easy to illustrate. Apart from facilitating the description of many-body contributions to the potential another important reason to use symmetry functions depending on all neighboring atoms is to obtain a constant number of function values independent of the number of atoms in the cutoff spheres. It is neither possible nor desirable to use a separate symmetry function for each neighboring atom, like a distance, because then the number of input nodes of the atomic NNs would change with the number of atoms in the cutoff spheres and the use of a single atomic NN per element would be impossible.

The most basic radial symmetry function describing the environment of atom i is simply a sum of the cutoff functions for all neighboring atoms j inside the cutoff sphere,

$$G_i^1 = \sum_{j=1}^{N_{\mathrm{atom}}} f_{\mathrm{c}}\left(R_{ij}\right). \tag{8}$$

Consequently, a single number is obtained, which can be interpreted as a coordination number up to the cutoff radius. This single number is of course not sufficient to describe the radial arrangement of the neighboring atoms. This can be achieved by using a set of these functions with different spatial extensions given by different cutoff radii as plotted in Fig. 5a. Still, this function type needs to be used with care as too short cutoff radii can give rise to artefacts in the forces close to the cutoff radius. A detailed discussion of this aspect can be found elsewhere[74].

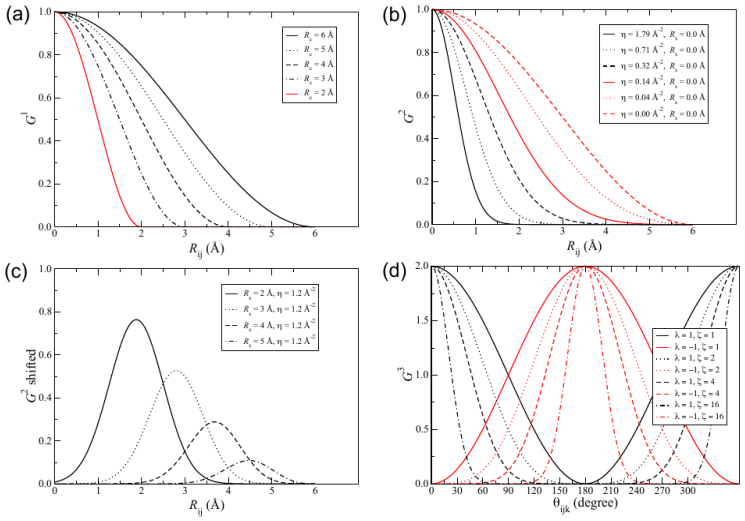


Figure 5: Plot of the radial symmetry functions G^1 and G^2 (panels (a)-(c)) and of the angular part of the angular symmetry function G^3 (panel (d)). In (b) and (c) the cutoff radius has been set to 6 Å.

A better choice to describe the radial arrangement of atoms in the cutoff sphere is a sum of products of Gaussians and the cutoff function,

$$G_i^2 = \sum_{j=1}^{N_{\mathrm{atom}}} e^{-\eta(R_{ij}-R_{\mathrm{s}})^2} \cdot f_{\mathrm{c}}\left(R_{ij}\right). \tag{9}$$

Here, for all functions the same large cutoff should be used to obtain reliable forces for all interatomic distances. Instead, the effective radial extension of the symmetry functions can be controlled by the parameter η, which defines the width of the Gaussians (cf. Fig. 5b). In this way, a strict decay to exactly zero in value and slope at too short radii can be avoided while still different radii can be probed by the symmetry functions using different values of η[74]. For non-zero values of the shifting parameter R_{s} the center of the Gaussians can also be displaced to certain radii resulting in spherical shells of a width given by η as shown in Fig. 5c to improve the sensitivity of the symmetry functions at specific radii.

Describing the radial distribution of neighboring atoms is not sufficient to obtain a suitable structural fingerprint of the atomic environments. This can be achieved by using additional angular functions depending on the angle θ_{ijk}, which is centered at the central atom i and is enclosed by the two interatomic distances R_{ij} and R_{ik}. As the potential is periodic with respect to this angle, instead of θ_{ijk} its cosine is used. A suitable angular symmetry function can then be defined as a sum over all cosines with repect to any possible pair of neighbors j and k, which are multiplied by Gaussians of the three interatomic distances in the triple of atoms and the respective cutoff functions,

$$\begin{aligned} G_i^3 = 2^{1-\zeta} \sum_{j\neq i} \sum_{k\neq i,j} \Big[& \left(1+\lambda\cdot\cos\theta_{ijk}\right)^{\zeta} \cdot e^{-\eta\left(R_{ij}^2+R_{ik}^2+R_{jk}^2\right)} \\ & \cdot f_{\mathrm{c}}\left(R_{ij}\right) \cdot f_{\mathrm{c}}\left(R_{ik}\right) \cdot f_{\mathrm{c}}\left(R_{jk}\right) \Big] \end{aligned} \tag{10}$$

This functional form ensures that G_i^3 approaches zero if the distance between any two atoms in the triple becomes larger than the cutoff radius, as in this case the triple is not completely included within the cutoff sphere of all three atoms. Again, the width of the Gaussians η needs to be specified to take into account that the importance of the angle depends on the atomic separations. The distribution of angles can be probed by using different exponents ζ as shown for the plot of the angular part of G_i^3 in Fig. 5d, and the normalization factor $2^{1-\zeta}$ ensures that the range of values is independent of the actual choice of exponent. The parameter λ, which can have values of $+1$ or -1 can be used to invert the shape of the cosine function to obtain a good description at different values of θ_{ijk}. While for $\lambda = 1$ the maxima of the cosine terms are at $\theta_{ijk} = 0°$, for $\lambda = -1$ they are located at $\theta_{ijk} = 180°$.

An alternative angular function is given by

$$G_i^4 = 2^{1-\zeta} \sum_{j\neq i} \sum_{k\neq i,j} \left[\left(1+\lambda\cdot\cos\theta_{ijk}\right)^{\zeta} \cdot e^{-\eta\left(R_{ij}^2+R_{ik}^2\right)} \cdot f_{\mathrm{c}}\left(R_{ij}\right) \cdot f_{\mathrm{c}}\left(R_{ik}\right) \right]. \tag{11}$$

Here, the Gaussians and cutoff functions with respect to R_{jk} have been omitted. This modification has two consequences. First, the numerical values of G_i^4 are generally larger than of G_i^3, since $e^{-\eta R_{jk}^2}$ and $f_c(R_{jk})$ are both lower than one, making this function more useful for larger atomic separations. Second, it is possible to take into account atomic triples

in which j and k are both in the cutoff sphere of atom i but separated by more than R_c. As the maximum possible distance between j and k inside the cutoff sphere is $2 \cdot R_c$, a significantly larger number of angles is thus included in the summation resulting in very different values compared to angular function G_i^3. Further types of radial and angular symmetry functions and a discussion of their properties can be found elsewhere[74]. The ability of these symmetry functions to describe and distinguish different atomic environments can also be used for the classification of structures, as has been demonstrated by Geiger and Dellago[104].

Several comments should be made on the definition of the symmetry function set. First, the parameters R_c, η, R_s, ζ, and λ define the spatial shape of the symmetry functions. Consequently, they are not optimized but remain fixed during the determination of the weight parameters of the NN. In this sense they are similar to basis functions in electronic structure calculations. Second, as the NN weights have been obtained for a specific set of input symmetry functions, these functions need to be known when applying the NNP in simulations. Consequently, the symmetry functions can be considered as a part of the NNP. Finally, the number of symmetry functions grows rapidly with the number of elements in the system. For instance, for each value of η there is one function of type G^2 for each element in the neighborhood and for each element of the central atom. As the atomic NNs are separated and independent from each other, the different possible central elements are not relevant. Still, the combinatorial increase of the radial and in particular of the angular functions for a given central atom is currently limiting the number of elements that can be included in high-dimensional NNPs. It has also been demonstrated, that high-dimensional NNPs can be constructed using environment-dependent atom-pair contributions to construct the potential-energy[105], but this method is compuationally more demanding than the atom-based approach, since the number of possible pairs is substantially larger.

4.3 Construction of Symmetry Function Sets

The choice of symmetry functions is a very important step for the construction of high-dimensional NNPs, as a reliable distinction of different structures is mandatory. If this distinction is not possible, e.g. if too few symmetry functions are used, different reference structures with differing energies and forces give rise to the same set of symmetry function values. Consequently, these structures are identical for the NN, the PES is not defined as a unique function of the atomic configuration, and there may be different energies for "the same structure", which results in poor fits.

While the construction of the symmetry function sets still is to some extent empirical, there is a number of analyses that can be carried out to investigate if a vector of symmetry functions is appropriate for the construction of a NNP:

- The symmetry functions should cover the configuration space in an unbiased way. A reasonable initial choice is to use, e.g., an equidistant set of radial functions as shown in Fig. 5b. The spatial extension of the function with the smallest effective range should be selected based on the shortest interatomic distances present in the data set. It must be avoided to include symmetry functions, which have a value of zero for all atoms and as a rule of thumb the turning point of the Gaussian with the largest η value should correspond to the shortst bond for the respective neighboring element.

Consequently, for different neighboring elements, the symmetry function parameters may differ.

- For each symmetry function the range of values present in the data set should be analyzed. If the range of values, i.e. the difference between the smallest and largest function value, is too small, the symmetry function is not contributing to the distinction of different structures and should be removed from the symmetry function set. Further, also the normalized range of values obtained through dividing by the standard deviation of the function values should be reasonable to ensure that the range of values is not dominated by a few outliers. If the difference between the symmetry function values is too small, the fits might become numerically unstable as the NN will try to assign substantial energy changes to tiny differences in the symmetry functions.

- In order to test if the symmetry functions allow to distinguish different structures, the data set should be investigated for contradictory data. In particular the atomic forces represent a valuable test case, as they provide local atom-specific information about the PES. If, for example, the magnitude of the forces is very different for two atoms, they must have a different chemical environment, and consequently also the symmetry function vectors must be different. If the data set is searched systematically for pairs of atoms having similar symmetry function vectors but experiencing different forces, as discussed in Ref.[74], inappropriate symmetry function sets can be identified and augmented by further functions until different atomic environments can be distinguished.

- The set of symmetry functions should be kept as small as possible to increase the efficiency of the calculation of the NN energy output. This can be investigated by determining the correlation between the values of a given symmetry function for all atoms in the reference set. If there is a high correlation, the symmetry functions are (close to) linearly dependent and essentially no additional information is provided by using both instead of only one of these symmetry functions. This analysis can also be used to construct symmetry functions in a systematic way. For instance, to find the η values for a set of G^2 functions, first a function with $\eta = 0$ is constructed, which has the largest possible spatial extension. For the next function, the η value is increased, i.e. the Gaussian is contracted, until the correlation of the obtained function values to the values of the first function decreases below a certain threshold value, e.g. 90 %. Then, the next function is added and η is further increased until the correlation to the other two functions is below this threshold an so forth. The set of G^2 functions is then complete, if an η value is reached, for which the range of values of the new symmetry function approaches zero.

All these tests for identifying a suitable set of symmetry functions depend on the specific composition of the available reference data set. Consequently, as the data set is typically increased step by step during the construction of a potential, several subsequent adjustments and refinements of the symmetry function set will be required. Usually it is not possible to use a large set of symmetry functions initially, since the small reference sets in the early stages of the construction of the NNP often only cover a part of the relevant configuration space. Functions, which enable to distinguish structural feature not present

in this set may thus not have a sufficiently wide range of values. On the other hand, the small symmetry function sets used for small initial data sets will not allow to distinguish data entering this set at a later stage. Consequently, an extension of the set of symmetry functions will be required while the reference set is increasing.

4.4 Preconditioning the Neural Network

In order to obtain a NNP, which is able to represent a given reference set with small errors, the fitting capabilities of NNs have to be exploited in an efficient way. This can be facilitated by preconditioning the atomic NNs in two different ways: concerning the input symmetry function values as well as concerning the initial values of the weight parameters.

By construction, the individual symmetry functions can have very different ranges of values, and a symmetry function with very large absolute values will have a siginificant impact on the value of the nodes in the first hidden layer, while symmetry functions with small absolute values will only play a minor role. The importance of the different symmetry functions can be balanced by rescaling the range of values for each symmetry function to the same interval[61], e.g. $[-1, 1]$, by applying

$$G_i^{\mathrm{scaled}} = \frac{2\left(G_i - G_{i,\mathrm{min}}\right)}{G_{i,\mathrm{max}} - G_{i,\mathrm{min}}} - 1, \tag{12}$$

where $G_{i,\mathrm{min}}$ is the smallest value of function G_i and $G_{i,\mathrm{max}}$ its largest value in the data set. It is often useful to center the range of values around 0, as this is the center of the non-linear regions of most activation functions. Alternatively, this can also be achieved by shifting the center of mass of all symmetry function values to zero. In addition to rescaling the symmetry functions, it is also possible to normalize the linear combinations at each node in the NN by dividing the sum by the number of nodes in the previous layer. This procedure has the advantage that it is applicable to arbitrary hidden layers and does not only refer to the first hidden layer. Further, even if the symmetry functions have been scaled, still the linear combination of symmetry function values can result in large numbers depending on the size of the NN, which is unfavorable as the subsequently applied activation functions are not used in their non-linear region but in their saturation part. Also this problem can be solved by normalizing the linear combination before applying the activation functions.

Apart from preconditioning the symmetry function values, also the initial choice of the weight parameters is very important. They can be chosen simply as random numbers, but a decision has to be made on the type of distribution as well as on the range of values. Further, there is a number of procedures that have been proposed in the literature to determine an optimum set of initial weight parameters, like the scheme of Nguyen and Widrow[106].

Finally, the initial errors of the training points can also be strongly reduced by preconditioning the weights. This primarily concerns the standard deviation and the center of mass of the output energies. The center of mass of the initially random energy outputs of the NNP can be aligned with the center of the target energies by adjusting the bias weight of the output node, which corresponds to a simple shift of the average output energy. The standard deviation of the NN output energies and of the energies in the training set can be matched by modifying the connecting weights between the last hidden layer and the output layer. By this simple procedure, the average initial errors of the NNP can often be reduced by one or two orders of magnitude.

4.5 Symmetry Functions and Forces

The availability of forces is crucial for many applications of NNPs like MD simulations and geometry optimizations. Since NNPs have well-defined functional forms, analytic derivatives are readily available by applying the chain rule to take into account the initial transformation of the atomic Cartesian coordinates onto the many-body symmetry functions. In general, the force with respect to some atomic coordinate α is the sum of the short range force $F_{\alpha,\mathrm{s}}$ and the electrostatic force $F_{\alpha,\mathrm{elec}}$,

$$F_\alpha = F_{\alpha,\mathrm{s}} + F_{\alpha,\mathrm{elec}} = -\frac{\partial E_\mathrm{s}}{\partial \alpha} - \frac{\partial E_\mathrm{elec}}{\partial \alpha}. \tag{13}$$

The short-range component of the force is then given by

$$\begin{aligned} F_{\alpha,\mathrm{s}} &= -\frac{\partial E_\mathrm{s}}{\partial \alpha} = -\sum_{j=1}^{N_\mathrm{atom}} \frac{\partial E_j}{\partial \alpha} \\ &= -\sum_{j=1}^{N_\mathrm{atom}} \sum_{\mu=1}^{N_{\mathrm{sym},j}} \frac{\partial E_j}{\partial G_{j\mu}} \cdot \frac{\partial G_{j\mu}}{\partial \alpha}, \end{aligned} \tag{14}$$

where $N_{\mathrm{sym},j}$ is the number of symmetry functions of atom j. The first term in the product is given by the architecture of the NN and contains also the weight parameters, while the second term depends on the definition of the symmetry functions.

The electrostatic force contribution can be calculated by

$$\begin{aligned} F_{\alpha,\mathrm{elec}} &= -\frac{\partial}{\partial \alpha}\frac{1}{2}\sum_{i=1}^{N_\mathrm{atom}} \sum_{j=1,j\neq i}^{N_\mathrm{atom}} \frac{Q_i Q_j}{R_{ij}} \\ &= -\sum_{i=1}^{N_\mathrm{atom}} \sum_{j=1,j\neq i}^{N_\mathrm{atom}} \frac{1}{2R_{ij}^2}\left[\frac{\partial Q_i}{\partial \alpha} Q_j R_{ij} + Q_i \frac{\partial Q_j}{\partial \alpha} R_{ij} - Q_i Q_j \frac{\partial R_{ij}}{\partial \alpha}\right]. \end{aligned} \tag{15}$$

Here, it is important to note that in addition to the usual derivative of Coulombs law that is used in many force fields employing fixed atomic charges there are two additional terms taking into account the environment-dependence of the atomic charges Q_i and Q_j. After some rearrangements, regrouping of indices and considering the transformation of the Cartesian coordinates to symmetry functions we finally obtain

$$F_{\alpha,\mathrm{elec}} = \sum_{j=1}^{N_\mathrm{atom}} \sum_{i=1,i\neq j}^{N_\mathrm{atom}} \frac{Q_i}{R_{ij}} \cdot \left[\frac{1}{2}\frac{Q_j}{R_{ij}}\frac{\partial R_{ij}}{\partial \alpha} - \sum_{k=1}^{N_{\mathrm{sym},j}} \frac{\partial Q_j}{\partial G_{jk}}\frac{\partial G_{jk}}{\partial \alpha}\right]. \tag{16}$$

In principle, also higher derivatives can be calculated, but due to the complex nested functional form of the NN this can become rather complicated. Also other physical properties including gradients like the stress tensor of solids are accessible[74].

It should be mentioned that the force acting on an atom depends on the positions of the atoms being as far as $2 \cdot R_\mathrm{c}$ away. At first glance, this seems odd, because the atomic energy contribution is determined only by the atoms inside the cutoff sphere. Still, this is fully consistent with the NN total energy expression, as the force with respect to a coordinate α of an atom i is the derivative of the energies E_j of all atoms j inside the cutoff sphere of i. In turn, the E_j depend on the positions of all atoms in their respective environments,

and these atoms can be up to $2 \cdot R_c$ away from atom i. As the positions of all atoms in the environment of j determine E_j, even an atom k being outside the cutoff sphere of atom i affects the contribution of atom i to E_j. Still, in practice this large effective range of the forces is not very important, since the effective physical range of the atomic interactions is usually much shorter than twice $R_c \approx 12 - 20$ Å.

5 Training Neural Networks

5.1 Selecting the Training Data

In contrast to conventional physical potentials NNs have a very flexible functional form without a physical origin. While this is the reason for the high accuracy that can be obtained in fitting the reference set, it can also give rise to large errors if NNPs are used to predict the energies for structures, which are very different from the configurations included in the training set.

There are two situations, when this can happen. First, NNs have very limited extrapolation capabilities, i.e., they usually fail outside the range of input values spanned by the training data. This concerns even very basic properties of the PES like the strong repulsion of atoms at very short distances, which needs to be learned from example structures in the data set. It is often found in early stages of the NNP construction that structures collapse and adopt unphysically short interatomic distances or "explode". In general these extrapolation cases are easy to identify. This is done by comparing the symmetry function vector of each atom in the structure with the minimum and maximum values of each symmetry function in the training data for the respective element. If a coordinate value is outside the range of validity of the potential, a warning can be issued and the simulation can be stopped. While this should not happen in the final application of the completed NNP, this warning turns out to be very useful during the construction of the NNP as it is possible to search systematically for such extrapolating structures to extend and improve the reference set.

The other situation is more difficult to identify. Here, the energy is requested for an atomic configuration, which is within the range of validity of the symmetry functions, but still the structure is located in a part of the configuration space that is not well represented in the training set. Uncontrolled errors can be present for such NN energies and forces. In case they are still in the expected order of magnitude, this is very difficult to detect. The reason for these errors is the high flexiblity of NNs, which can cause strong artifacts in between sparsely distributed training points. This is the well-known "overfitting" problem, which is essentially a much better representation of the points in the training set than of the structures in between.

If the reference data set covers all parts of configuration space but is just too sparse, there is a simple recipe to detect overfitting, which is called the "early stopping" method. Here, the available reference data set is split into the training set, which is used for the optimization of the weight parameters, and an independent test or validation set, whose error is monitored during the fit but which is not used in the weight optimization. If the errors of the training and the test set are similar, the NNP has good generalization properties and it is applicable also to structures not included in the training set. If, on the other hand, the error of the test data is significantly larger than the error of the training set, overfitting

is present and more training data is required. The typical evolution of a fit exhibiting overfitting is shown in Fig. 6. In the first iterations the errors of both the training and the

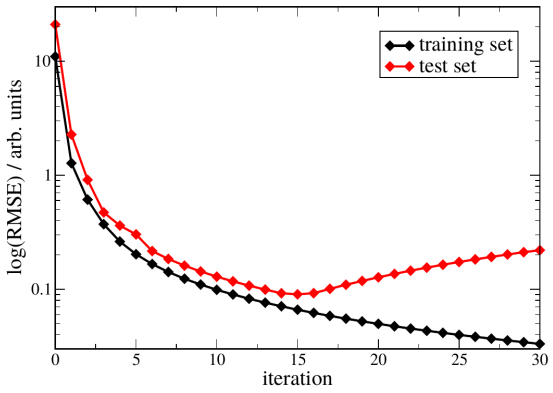


Figure 6: Illustration of the early stopping method. Initially, the errors of the training and the test set decrease as the neural network learns the overall shape of the potential-energy surface. Then, the test error reaches a minimum and starts to increase slowly, which labels the onset of overfitting.

test set decrease, because the NN learns the overall shape of the PES. Then, the error of the test set reaches a minimum and starts to increase again slowly. At this stage, the NN is learning the fine details of the training points at the expense of a larger error of the test set, which is not visible to the NN optimization algorithm. Therefore, the best set of weight parameters corresponds to the epoch with the smallest test set error.

Unfortunately, the early stopping method is not generally applicable and is particularly problematic in case of high-dimensional PESs. Problems can arise, if there are parts of the configuration space in which no reference data are present at all, since in this case failures of the NNP would not be detected by a high error of the test set. Still, there is a possibility to check and improve the NNP also in this situation without the need to carry out costly electronic structure calculations for many structures just to validate the potential, which is anyway impractical in a systematic way for high-dimensional PESs. This validation is done by first fitting several NNPs to the same training set. These NNPs must have different functional forms, which can be achieved e.g. by using atomic NNs with different architectures. Then, several NNPs of comparable quality are selected, which are apparently all equally applicable and accurate. Then a large number of structures is generated using one of these NNPs, e.g. by running NN-driven MD simulations. Afterwards, the obtained trajectory is recalculated using the other NNPs and the energies along the trajectory are compared as shown in Fig. 7. In panel (a) only a few structures representing the target function are used in the training and both NNPs agree well with the this function shown as black line close to these points. Far away from these points, however, the energy predictions of both NNPs are more or less random. This can be detected by comparing the predictions of several NNs. If there is a significant deviation between the predictions of the available NNPs for a given structure, then this structure is too far away from the training points, an electronic structure calculation should be carried out and it should be

included in the training set to refine the NNP. The result is shown schematically in panel (b) of Fig. 7. Here, the number of training points is increased while the region in which both NNPs deviate from each other is strongly reduced. After further improvements both NNPs agree well with the target function in panel (c) over the whole range. This multiple-fit validation should be carried out for the same simulation type and under the conditions of the intended applications of the NNP to ensure that the relevant part of configuration space is validated. As long as structures are found for which different fits having similar errors for the training set predict different energies or forces for some structures, the potential needs to be improved by adding further training structures to refine the values of the weight parameters. As soon as no problematic structures are found anymore, the NNP is reliable and the potential is ready for use. Still, ideally the final simulations should be carried out independently by different NNPs for validation.

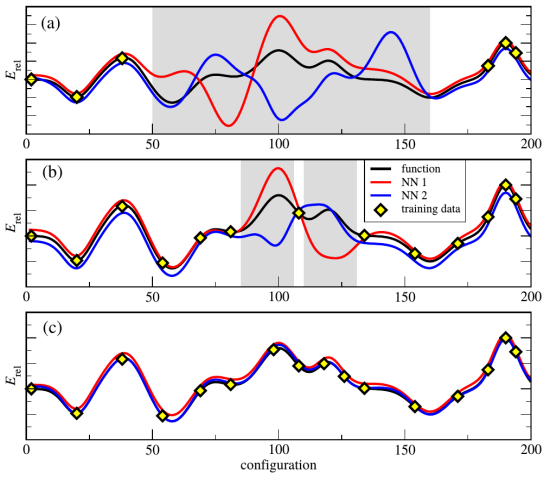


Figure 7: Illustration of the systematic improvement of the training set using the multiple-fit procedure. In panel (a) only a few training point of the target function (black line) are available and in between these points different neural networks can predict very different energies E_{rel}. As more training points are added (panel (b)), the representation of the energy function becomes more reliable until a good fit is obtained over the full range in panel (c).

Typically, the resulting reference data sets for the construction of NNPs are very large, starting from about 1000 structures to get an approximate PES for low-dimensional systems containing only a few atoms, up to several tens of thousands of structures for high-dimensional condensed systems. This poses some restrictions on the possible choice of reference electronic structure method. To date, in most cases DFT has been employed, in particular for large systems, but for small molecules also a variety of higher-level wave function-based methods has been used. The reference method is very important for the scientific problem to be solved, as NNPs cannot provide better results than the underlying electronic structure method.

Concerning the composition of the training sets for high-dimensional NNPs, it is possible to combine different types of structures including periodic and non-periodic systems

and configurations with different numbers of atoms. Due to the reduced computational costs it is often helpful to start the NNP development with a set of structures for very small systems like thermally distorted crystal unit cells to get a first rough estimate for basic features of the PES like prefered interatomic distances and lattice constants. Then, as soon as NN-based simulations of these small systems work reliably, the size of the system can be increased step by step until the limit of what can be addressed by the reference electronic structure has been reached. Still, often it is required to validate the potential also for systems that are much larger. In this case the effective reduction of dimensionality by the symmetry function cutoff can be exploited by cutting clusters centered at atoms that are not reliably represented by the NNP. These clusters, which can be identified e.g. by the multiple NNP method and investigating the forces as local probes of the PES, can then be recalculated by the reference method to include the environments of these atoms in the training set. In principle, it is sufficient to include the configurations that can be realized within the cutoff spheres in the training set to obtain a NNP that can be applied to systems of arbitrary size. Finally, electronic structure calculations for clusters can also be employed to directly validate the NN PES for very large systems using the forces. If the clusters are sufficiently large, which can be checked in convergence tests, the forces acting on the central atoms obtained in electronic structure calculations should be very similar to the NN forces for the full system, and this has indeed been found for several systems[77,83].

In summary, the determination of the reference set can be done in a self-consistent way according to the following procedure:

1. Select an electronic structure method, which is suffiently fast to carry out the required number of reference calculations and which is accurate enough to describe the physical properties of interest reliably.

2. Define a first set of structures and determine the reference energies and forces.

3. Construct a first preliminary NNP.

4. Carry out simulations using this NNP to find structures, which give rise to extrapolation warnings or unphysical geometries.

5. Determine the electronic structure energies and forces of these structures, include them in the training set and improve the NNP.

6. Improve the NNP systematically and self-consistently by running NNP-based simulations to find missing structures using the multiple-NNP method. Carry out electronic structure calculations for these structures, refine the fit and start again running extended simulations at the conditions of the intended applications.

7. As soon as no further problematic structures can be identified, the NNP is ready for use.

There is some flexibility in the choice of the first set of structures. In a “puristic” approach it is possible start with random structures to obtain an unbiased initial data set. This is more demanding as random structures typically have a high energy and will not be relevant in the final converged data set. Still, starting with such structures it is possible to improve the potential step by step until the correct structures are found. We have tested this for boron.

Finally, after several iterative improvements employing metadynamics simulations[107], we have obtained a potential suitable to identify the correct crystal structure of α-boron[108]. In a more "pragmatic" approach, which employs some knowledge about the system, the number of electronic structure calculations can be reduced by starting from reasonable atomic configurations. Further, also other types of potentials like classical force fields could be used to run the first simulations. In the later stages of the iterative refinement of the potential it is still crucial to run simulations using the NNP itself to find problematic configurations like "holes" in the PES.

The large number of training points is one of the main remaining challenges of constructing NNPs. Still, there is some guidance how an excessive growth of the training set can be avoided. First of all, only the part of configuration space, which is accessible in MD simulations is relevant and needs to be mapped. Consequently, calculating the data on a regular grid is not required and is unfeasible anyway in high-dimensional systems as the number of points grows exponentially with the number of dimensions. Also using a metric to determine positions of candidate structures is not a promising approach as the "distance" between structures depends on the choice of coordinates and can be very different in real space and in the symmetry function space. Further this approach is sensitive to the preprocessing steps of the symmetry functions like scaling and centering. Further, as much information as possible should be extracted from each electronic structure calculation. While there is just one total energy value, there are $3 \cdot N_{\mathrm{atom}}$ force components containing valuable information about the local topology of the PES. This can be used to reduce the training set size significantly. It has also been proposed to employ the symmetry, e.g. of single crystal surfaces[50], but this is possible only in special applications as in general there is no symmetry left in MD simulations at finite temperatures.

5.2 Determination of the Weight Parameters

Once a set of reference data from electronic structure calculations is available, the central step in constructing a NNP is the determination of the numerical values of the weight parameters, which enable to reproduce these data as accurately as possible. In "supervised learning", which is the standard procedure in the training of NNPs, the output of the NN for each structure is compared to the known "true" answer from the reference calculations. In general this corresponds to a minimization of an error function Γ, which is given as the sum of squared errors of the N_{struct} individual members of the training set,

$$\Gamma = \frac{1}{N_{\mathrm{struct}}} \sum_{i=1}^{N_{\mathrm{struct}}} \left(E^i_{\mathrm{NN}} - E^i_{\mathrm{Ref}}\right)^2 . \tag{17}$$

If also the forces are used for the weight optimization, a modified error function,

$$\Gamma = \frac{1}{N_{\mathrm{struct}}} \sum_{i=1}^{N_{\mathrm{struct}}} \left[(E^i_{\mathrm{NN}} - E^i_{\mathrm{Ref}})^2 + \frac{\beta}{3N^i_{\mathrm{atom}}} \sum_{j=1}^{3N^i_{\mathrm{atom}}} (F^i_{j\mathrm{NN}} - F^i_{j\mathrm{Ref}})^2 \right], \tag{18}$$

can be used, which contains a loop over all X, Y, and Z force components of all atoms. The relative influence of the energies and forces can be balanced by the parameter β, as the number of force components is much larger than the number of energies. The value of β can be set to the inverse of the number of force components for each structure to take

into account structures with different numbers of atoms, or to a predefined fixed value. Alternatively, we have found it useful to first update the weights using the error of the energy, followed by one update per force component. After each of these force updates the energy update is repeated to stress the importance of a correct representation of the energy.

The optimization of the weight parameters is done iteratively, and one iteration, which is often called "epoch" in the context of NNs, corresponds to a cycle through the full data in the training set. In "batch learning", which is also called offline learning, the weights are updated only once per iteration, while in "online learning" there is one weight update after the presentation of each piece of information, like energies or force components. While online learning is computationally more demanding, in this procedure the fit is less likely to get trapped in local minima. This is an important advantage since there are typically several thousand parameters, and there is no hope to find the global minimum in this very high-dimensional optimization space. Still, usually many sufficiently accurate local minima are found, which provide a reliable description of the PES.

A wide range of optimization algorithms can be used to determine the weight parameters, and the standard procedure is to use gradient-based approaches. The most basic method is called "backpropagation" in the NN community[109], which refers to the recursive calculation of the derivatives of the error function with respect to the weights starting from the output layer and proceeding to the input layer in the inverse order of the calculation of the NN output. In essence, the backpropagation algorithm corresponds to a steepest descent optimization of each weight w, which can be a connecting weight a or a bias weight b. For an iteration $t+1$ the updated weights are obtained according to

$$w(t+1) = w(t) - \eta \cdot \frac{\partial E}{\partial w(t)}. \tag{19}$$

Here, η is the learning rate, which can also be adapted during the progress of the fit. More advanced optimization algorithms that are frequently used for the determination of NN weights are the Levenberg-Marquardt algorithm[110] and in particular the global extended Kalman filter[111,112,93,113], which has originally been developed as a method to derive parameters from nonlinear measurements generated in real-time. A discussion of the specific details of these optimization algorithms is beyond the scope of this lecture and can be found in the provided references.

Regardless of the specific choice of optimization algorithm, for each energy the derivatives of the error function with respect to the connecting weights and the bias weights need to be determined for each atomic energy, as the reference energy is a constant and thus independent of the weights. Similar derivatives also need to be calculated for each force component. This has to be repeated after each weight update since the derivatives with respect to the weights depend also on their numerical values.

During the optimization the quality of the fit is measured by determining the root mean squared error (RMSE) of the energies and forces in the training and the test set in each iteration using

$$RMSE(E) = \sqrt{\frac{1}{N_{\mathrm{struct}}} \sum_{i=1}^{N_{\mathrm{struct}}} \left(E^i_{\mathrm{NN}} - E^i_{\mathrm{Ref}}\right)^2} \tag{20}$$

and

$$RMSE(F) = \sqrt{\frac{1}{N_{\text{struct}}} \sum_{i=1}^{N_{\text{struct}}} \sum_{j=1}^{3N_{\text{atom}}^i} \frac{1}{3N_{\text{atom}}^i} \left(F_{j,\text{NN}}^i - F_{j,\text{Ref}}^i\right)^2}. \tag{21}$$

In particular in case of large data sets including very different systems like cluster and bulk data, a global RMSE is not very helpful in understanding how accurate a fit will be in which situation. For a more detailed analysis the data set should be split in meaningful groups, like subsets of structures with the same number of atoms, and the error of each group should be determined for a given fit. In order to obtain a meaningful measure of the accuracy of the fit, the RMSE should also be normalized per atom to enable a comparison of different types of structures. Additionally, also points with particularly large errors should be identified and analysed to understand the reason why they are not well-represented. A possible reason could be for instance an insufficient set of symmetry functions or problems in the underlying reference data. If specific parts of the configuration space, like transition states and local minima are of special interest, it is also possible to increase their impact on the fit[29,50], but in many situations this distinction of the PES in more and less important parts is not straightforward.

Apart from the selection of the optimization algorithm, a number of additional choices have to be made to construct a specific NNP. First of all, the architecture of the atomic NNs has to be determined. This is still an empirical component of the NNP development, since the most efficient way to identify a suitable number of hidden layers and nodes per layer is simply to carry out a number of fits and to select the one with the lowest errors of the energies and forces in the test set. In general, the architecture of the NN determines its flexibility. Consequently, if the NN is too small, some features of the PES may not be resolved as shown in Fig 8a. This is visible in the RMSEs as a notable error in the training and the test set. If the flexbility of the NN is increased by using more nodes and possibly an additional hidden layer, the representation of the PES will improve, as demonstrated in Fig. 8b, resulting in lower errors of the training and the test data. If the NN is too large, it becomes too flexible and overfitting can occur, which is illustrated in Fig. 8c and can be identified by an increased error of the test set data compared to the error of the training data. The tendency for overfitting can be reduced significantly, if also forces are used for the training, because a correct representation of the gradient of the PES close to the training points will improve the description of neighboring structures.

Other choices in the fitting process can have a strong influence on the CPU time requirements. In particular if also forces are used for the determination of the weight parameters, very large data sets often containing millions of pieces of information have to be processed in each iteration. This comprises first the determination of the error of each energy and force component, followed by the calculation of the derivative of this error with respect to each weight parameter and finally the update of the weights, which e.g. in case of the Kalman filter involves demanding matrix operations. Consequently, it should be avoided to perform unnecessary weight updates, e.g. if an energy or force is already well represented. This can be done be introducing error thresholds and only if the error of an energy or force is above this threshold a weight update will be carried out. This threshold is usually coupled to the RMSE of the present iteration and therefore decreases along with the overall error of the NNP. Another possibility to reduce the fitting effort is to use only a random

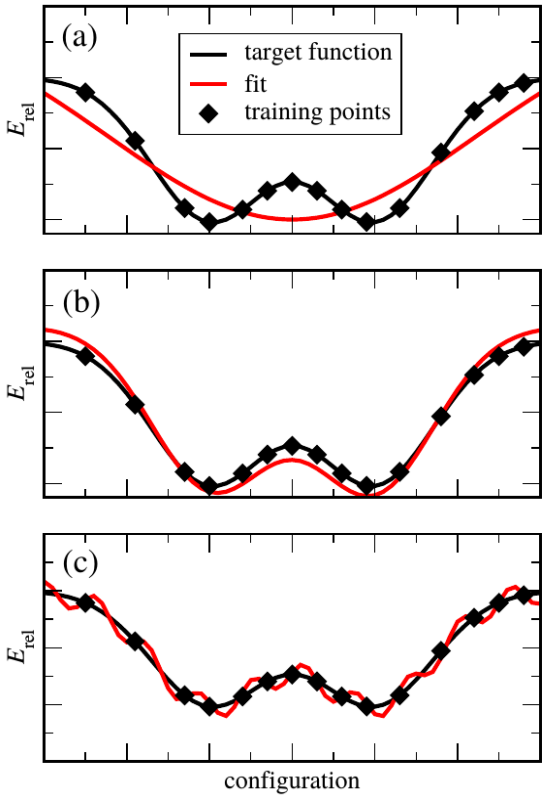


Figure 8: Influence of the flexibility of the neural network (NN) on the fit quality. In panel (a) a very small NN is used, which is not flexible enough to resolve all features of the energy surface. In (b) the shape of the energy surface is well represented by using a large NN, while in (c) overfitting is present as can be seen by the presence of artificial wiggles in between the training points.

subset of the forces in each iteration or to average the weight derivatives for several energies and/or forces and to perform a joint weight update after a certain number of weight derivatives has been accumulated.

Once the weights have been determined, the NNP is ready for use, and it is not relevant for the application how the weights have been obtained. Still, it is important to be aware that the "NNP" not only consists of the values of the weight parameters. Additionally, information about the architecture of the NN, the type of activation functions, the types and parameters of the symmetry functions, and possibly also scaling information for the preconditioning of the NN input nodes are required to apply a NNP in simulations. Ideally, the same NN software should be used for the simulations that has also been used for the construction of the NNP, since many subtle details of the implementation may complicate the transfer of this information from one code to another.

6 Discussion

Having reviewed the structure of high-dimensional NNPs, the functional form of the symmetry functions and the determination of the training sets as well as of the weight parameters, some comments should be made concerning the general applicability and performance of high-dimensional NNPs. Like other machine learning methods, NNPs can be numerically very accurate and the typical RMSEs for the high-dimensional NNP method presented here are less than 10 meV/atom for potential-energies and between 50 and 200 meV/Bohr for the atomic forces. For energy differences, which are most important for the investigation of chemical and physical processes, the error can be even about one order of magnitude smaller. This has the consequence that replacing electronic structure methods like DFT in MD simulations by NNPs does not give rise to significant additional errors, which is crucial as substantial uncontrolled errors arising from the fit would not be acceptable. On the other hand, the low RMSEs of NNPs pose stringent conditions on the convergence level of the training data, as any noise in the reference calculations would complicate the optimization of the weight parameters. As a rule of thumb the convergence error of the energies and forces should be one order of magnitude smaller than the intended final RMSEs of the NNP.

Concerning the efficiency, NNPs are superior to any electronic structure method, and since there is one atomic NN per atom, the method scales linearly with the number of atoms irrespective of the system size. Further, it is straightforward to implement in parallel computer codes. On current desktop computers, depending on the complexity of the selected symmetry functions, the energies and forces of about 200 atoms can be calculated per second and per compute core. Still, this is substantially slower than the evaluation of basic classical force fields. NNPs, however, are no direct competitors of classical force fields due to their different applicability. NNPs are currently restricted to systems containing about three to four elements because of the rapidly increasing structural variety that would give rise to too many symmetry functions for systems consisting of more elements. This restriction is not present for force fields, which employ only very low-dimensional additive terms. NNPs, on the other hand, are clearly superior to force fields concerning the numerical accuracy and their ability to work even in case of most complex geometries. Further, they are able to describe the making and breaking of bonds, which is a severe problem for most force fields. Like electronic structure methods, the input of NNPs consists of a description of the chemical elements and the atomic positions, but no classification of atoms according to functional groups or the definition of bonds is required. In contrast to common belief, also the number of fitting parameters in classical force fields can be very large taking into account the interaction between all the different atom types and bonds.

A clear drawback of NNPs is the non-physical functional form, which does not allow for a physical interpretation. Still, this seems to be the price to be paid for obtaining potentials with a high numerical accuracy. Consequently, NNPs must be constructed and used with care. Since the functional form is not transparent to the user, NNPs should not be used as black box method but must be routinely tested using the methods described above.

There is still a number of open problems concerning the construction and application of high-dimensional NNPs. First, as the development of NNPs currently requires very large training sets, a major goal is to reduce the number of required reference data. One step in this direction is the use of forces apart from total energies only, which provide

a lot of additional information about the local shape of the PES. Nevertheless, there is certainly room for a better and more selective choice of reference structures, although still all relevant parts of configuration space need to be mapped, which is posing a boundary on the minimum number of structures that need to be included. The limitation of NNPs to a few elements can be partially overcome by sacrificing the full generality of NNPs. If, for instance, it is known that certain bonds do not participate in chemical processes, the NNPs does not need to be able to describe this event. Finally, also the fitting process itself can be very time consuming and most efficient online-learning algorithms like the Kalman filter cannot be parallelized well.

To date, NNPs have been applied successfully in different types of simulations like Monte Carlo[60,61], molecular dynamics[29], metadynamics[75] and replica exchange MD simulations[85], and it can be anticipated that the use of ML potentials and in particular of NNPs will rapidly increase in the near future.

7 Summary

In this lecture, a method to construct high-dimensional atomistic potentials employing artificial neural networks has been presented and discussed. In contrast to conventional, low-dimensional NNPs, the potential-energy of the system is not represented by a single feed-forward NN. Instead, there is one separate NN for each atom in the system providing the contribution of an atom to the total energy. These atomic energy contributions depend on the local chemical environment defined by a cutoff radius resulting in atomic spheres containing about 100 atoms. The positions of these atoms with respect to the central atom are described by a vector of many-body symmetry functions serving as input for the atomic NNs. In addtion to this short-range energy also long-range electrostatic energies can be included based on environment-dependent charges represented by another set of atomic NNs.

The method is applicable to a wide range of systems and very low errors with respect to the underlying electronic structure data can be obtained. The evaluation of the NNP is several orders of magnitude faster than DFT even for systems of moderate size and scales linearly with the number of atoms. Still, due to the very flexible functional form of NNs the construction of the PES requires care and the obtained NNP must be thoroughly validated. The determination of the training set, which usually consists of tens of thousands of structures, is computationally demanding. An important limitation of the method is its current restriction to a few elements, although a large number of atoms of each element can be present, because of the rapidly increasing structural variety in the atomic environments, which needs to be mapped by electronic structure calculations and has to be described by the symmetry functions.

Since NNPs are equally applicable to all types of bonding and even complex atomic environments, they are particularly useful for large-scale MD and Monte Carlo simulations of complex systems in materials science, at interfaces like the solid-liquid interface, and for studying solvation processes.

Acknowledgements

This work was supported by the Cluster of Excellence RESOLV (EXC 1069) funded by the Deutsche Forschungsgemeinschaft as well as by the Emmy Noether program (BE3264/3-1) and a Heisenberg fellowship (BE3264/6-1) of the DFG.

References

1. This lecture will also be published as a tutorial review titled "Constructing High-Dimensional Neural Network Potentials: A Tutorial Review" in a themed issue on "Combining Quantum Chemistry with Machine Learning" in the International Journal of Quantum Chemistry.
2. M. Born and R. Oppenheimer, Ann. Phys. **389**, 457–484, 1927.
3. R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471–2474, 1985.
4. D. Marx and J. Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*, Cambridge University Press, Cambridge, 2009.
5. R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford, 1989.
6. W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2nd edition, Wiley-VCH, 2001.
7. N. L. Allinger, Y. H. Yuh, and J.-H. Lii, J. Am. Chem. Soc. **111**, 8551–8566, 1989.
8. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, Jr., K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179–5197, 1995.
9. S. L. Mayo, B. D. Olafson, and W. A. Goddard III, J. Phys. Chem. **94**, 8897–8909, 1990.
10. A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard III, and W. M. Skiff, J. Am. Chem. Soc. **114**, 10024–10035, 1992.
11. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187–217, 1983.
12. F. H. Stillinger and T. A. Weber, Phys. Rev. B **31**, 5262–5271, 1985.
13. J. Tersoff, Phys. Rev. Lett. **56**, 632–635, 1986.
14. M. W. Finnis, Prog. Mater. Sci. **52**, 133–153, 2007.
15. A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard III, J. Phys. Chem. A **105**, 9396–9409, 2001.
16. C. A. Coulson, Proc. Roy. Soc. London **A169**, 413–428, 1939.
17. A. Brown, B. J. Braams, K. Christoffel, Z. Jin, and J. M. Bowman, J. Chem. Phys. **119**, 8790–8793, 2003.
18. J. M. Bowman, G. Czakó, and B. Fu, Phys. Chem. Chem. Phys. **13**, 8094, 2011.
19. J. Ischtwan and M. A. Collins, J. Chem. Phys. **100**, 8080–8088, 1994.
20. C. Crespos, M. A. Collins, E. Pijper, and G. J. Kroes, Chem. Phys. Lett. **376**, 566–575, 2003.
21. A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403, 2010.
22. A. P. Bartók, M. J. Gillan, F. R. Manby, and G. Csányi, Phys. Rev. B **88**, 054104, 2013.
23. M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, Phys. Rev. Lett. **108**, 058301, 2012.
24. K. Hansen, G. Montavon, F. Biegler, S. Fazli, M. Rupp, M. Scheffler, O. A. von Lilienfeld, A. Tkatchenko, and K.-R. Müller, J. Chem. Theory Comp. **9**, 3404—3419, 2013.
25. M. J. L. Mills, and P. L. A. Popelier, Theor. Chem. Acc. **131**, 1137, 2012.

26. T. Ishida and G. C. Schatz, Chem. Phys. Lett. **314**, 369–375, 1999.
27. R. Dawes, D. L. Thompson, Y. Guo, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **126**, 184108, 2007.
28. T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, J. Chem. Phys. **103**, 4129–4137, 1995.
29. S. Lorenz, A. Groß, and M. Scheffler, Chem. Phys. Lett. **395**, 210–215, 2004.
30. A. Vitek, M. Stacho, P. Kromer, and V. Snasel, in *International Conference on Intelligent Networking and Collaborative Systems (INCoS)* pp. 121-–126, 2013.
31. C. M. Handley and J. Behler, Eur. Phys. J. B **87**, 152, 2014.
32. C. M. Bishop, *Neural Networks for Pattern Recognition*; Oxford University Press, Oxford, 1996.
33. S. Haykin, *Neural Networks and Learning Machines*; 3rd International Edition; Prentice Hall: New York, 2009.
34. W. McCulloch and W. Pitts, Bull. Math. Biophys. **5**, 115–133, 1943.
35. J. Gasteiger and J. Zupan, Angew. Chem. Int. Ed. **32**, 503–527, 1993.
36. T. Kohonen, Neural Networks **1**, 3–16, 1988.
37. G. Cybenko, Math. Control Signals Systems **2**, 303–314, 1989.
38. K. Hornik, M. Stinchcombe, and H. White, Neural Networks **2**, 359–366, 1989.
39. F. Scarselli and A. V. Tsoi, Neural Networks **11**, 15–37, 1998.
40. J.-G. Attali and G. Pages, Neural Networks **10**, 1069–1081, 1997.
41. K. Hornik, Neural Networks **4**, 251–257, 1991.
42. J. Behler, J. Phys.: Condens. Matter **26**, 183001, 2014.
43. F. V. Prudente and J. J. Soares Neto, Chem. Phys. Lett. **287**, 585–589, 1998.
44. F. V. Prudente, P. H. Acioli, and J. J. Soares Neto, J. Chem. Phys. **109**, 8801–8808, 1998.
45. K. T. No, B. H. Chang, S. Y. Kim, M. S. Jhon, and H. A. Scheraga, Chem. Phys. Lett. **271**, 152–156, 1997.
46. H. M. Le, S. Huynh, and L. M. Raff, J. Chem. Phys. **131**, 014107, 2009.
47. H. M. Lee and L. M. Raff, J. Chem. Phys. **128**, 194310, 2008.
48. J. Ludwig and D. G. Vlachos, J. Chem. Phys. **127**, 154716, 2007.
49. J. Behler, S. Lorenz, and K. Reuter, J. Chem. Phys. **127**, 014705, 2007.
50. S. Lorenz, M. Scheffler, and A. Groß, Phys. Rev. B **73**, 115431, 2006.
51. J. Behler, B. Delley, S. Lorenz, K. Reuter, and M. Scheffler, Phys. Rev. Lett. **94**, 36104, 2005.
52. J. Behler, K. Reuter, and M. Scheffler, Phys. Rev. B **77**, 115421, 2008.
53. C. Carbogno, J. Behler, A. Groß, and K. Reuter, Phys. Rev. Lett. **101**, 096104, 2008.
54. C. Carbogno, J. Behler, K. Reuter, and A. Groß, Phys. Rev. B **81**, 035410, 2010.
55. I. Goikoetxea, J. Beltrán, J. Meyer, J. I. Juaristi, M. Alducin, and K. Reuter, New J. Phys. **14**, 013050, 2012.
56. I. Goikoetxea, J. Meyer, J. I. Juaristi, M. Alducin, and K. Reuter, Phys. Rev. Lett. **112**, 156101, 2014.
57. D. A. R. S. Latino, R. P. S. Fartaria, F. F. M. Freitas, J. Aires-de-Sousa, and F. M. S. S. Fernandes, J. Electroanal. Chem. **624**, 109–120, 2008.
58. S. Manzhos and K. Yamashita, Surf. Sci. **604**, 554–560, 2010.
59. T. Liu, B. Fu, and D. H. Zhang, Science China Chemistry **57**, 147–155, 2014.
60. K.-W. Cho, K. T. No, and H. A. Scheraga, J. Mol. Struct. **641**, 77–91, 2002.

61. H. Gassner, M. Probst, A. Lauenstein, and K. Hermansson, J. Phys. Chem. A **102**, 4596–4605, 1998.
62. C. M. Handley and P. L. A. Popelier, J. Phys. Chem. A **114**, 3371–3383, 2010.
63. J. Behler, Phys. Chem. Chem. Phys. **13**, 17930–17955, 2011.
64. S. Hobday, R. Smith and J. Belbruno, Modelling Simul. Mater. Sci. Eng. **7**, 397–412, 1999.
65. S. Hobday, R. Smith, and J. BelBruno, Nucl. Instr. Meth. Phys. Res. B **153**, 247–263, 1999.
66. A. Bholoa, S. D. Kenny, and R. Smith, Nucl. Instr. Meth. Phys. Res. B **255**, 1–7, 2007.
67. E. Sanville, A. Bholoa, R. Smith, and S. D. Kenny, J. Phys.: Condens. Matter **20**, 285219, 2008.
68. S. Manzhos and T. Carrington, Jr., J. Chem. Phys. **125**, 84109, 2006.
69. S. Manzhos and T. Carrington, Jr., J. Chem. Phys. **125**, 194105, 2006.
70. S. Manzhos and T. Carrington, Jr., J. Chem. Phys. **127**, 014103, 2007.
71. S. Manzhos and T. Carrington, Jr., J. Chem. Phys. **129**, 224104, 2008.
72. S. Manzhos, K. Yamashita, and T. Carrington, Jr., Comp. Phys. Comm. **180**, 2002–2012, 2009.
73. J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401, 2007.
74. J. Behler, J. Chem. Phys. **134**, 074106, 2011.
75. J. Behler, R. Martoňák, D. Donadio, and M. Parrinello, Phys. Rev. Lett. **100**, 185501, 2008.
76. J. Behler, R. Martoňák, D. Donadio, and M. Parrinello, Phys. Status Solidi B **245**, 2618–2629, 2008.
77. N. Artrith and J. Behler, Phys. Rev. B **85**, 045439, 2012.
78. R. Z. Khaliullin, H. Eshet, T. D. Kühne, J. Behler, and M. Parrinello, Phys. Rev. B **81**, 100103, 2010.
79. R. Z. Khaliullin, H. Eshet, T. D. Kühne, J. Behler, and M. Parrinello, Nature Materials **10**, 693–697, 2011.
80. H. Eshet, R. Z. Khaliullin, T. D. Kühne, J. Behler, and M. Parrinello, Phys. Rev. B **81**, 184107, 2010.
81. H. Eshet, R. Z. Khaliullin, T. D. Kühne, J. Behler, and M. Parrinello, Phys. Rev. Lett. **108**, 115701, 2012.
82. N. Artrith, T. Morawietz, and J. Behler, Phys. Rev. B **83**, 153101, 2011.
83. N. Artrith, B. Hiller, and J. Behler, Phys. Status Solidi B **250**, 1191–1203, 2013.
84. T. Morawietz, V. Sharma, and J. Behler, J. Chem. Phys. **136**, 064103, 2012.
85. T. Morawietz and J. Behler, J. Phys. Chem. A **117**, 7356–7366, 2013.
86. T. Morawietz and J. Behler, Z. Phys. Chem. **227**, 1559–1581, 2013.
87. S. Kondati Natarajan, T. Morawietz, and J. Behler, Phys. Chem. Chem. Phys. , in press, 2014. DOI: 10.1039/C4CP04751F.
88. G. C. Sosso, G. Miceli, S. Caravati, J. Behler, and M. Bernasconi, Phys. Rev. B **85**, 174103, 2012.
89. G. C. Sosso, D. Donadio, S. Caravati, J. Behler, M. Bernasconi, Phys. Rev. B **86**, 104301, 2012.
90. G. C. Sosso, J. Behler, and M. Bernasconi, Phys. Status Solidi B **249**, 1880–1885, 2012.

91. G. C. Sosso, G. Miceli, S. Caravati, F. Giberti, J. Behler, and M. Bernasconi, J. Phys. Chem. Lett. **4**, 4241—4246, 2013.
92. C. Muñoz-Caro and A. Niño, Computers Chem. **22**, 355–361, 1998.
93. J. B. Witkoskie and D. J. Doren, J. Chem. Theory Comput. **1**, 14–23, 2005.
94. P. P. Ewald, Ann. Phys. **64**, 253–287, 1921.
95. P. Salvador and I. Mayer, J. Chem. Phys. **126**, 234113, 2007.
96. I. Mayer, Phys. Chem. Chem. Phys. **8**, 4630–4646, 2006.
97. F. L. Hirshfeld, Theor. Chim. Acta **44**, 129–138, 1977.
98. S. Houlding, S. Y. Liem, and P. L. A. Popelier, Int. J. Quantum Chem. **107**, 2817–2827, 2007.
99. M. G. Darley, C. M. Handley, and P. L. A. Popelier, J. Chem. Theor. Comput. **4**, 1435–1448, 2008.
100. S. Grimme, WIREs Comput. Mol. Sci. **1**, 211, 2011.
101. S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104, 2010.
102. A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005, 2009.
103. B. J. Braams and J. M. Bowman, Int. Rev. Phys. Chem. **28**, 577–606, 2009.
104. P. Geiger and C. Dellago, J. Chem. Phys. **139**, 164105, 2013.
105. Jovan Jose K. V., N. Artrith, and J. Behler, J. Chem. Phys. **136**, 194111, 2012.
106. D. H. Nguyen and B. Widrow, IEEE Control Syst. Mag. **3**, 18–23, 1990.
107. A. Laio and M. Parrinello, Proc. Nat. Acad. Sci. USA **99**, 12562–12566, 2002.
108. T. Morawietz and J. Behler, *Unpublished results.*
109. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Nature **323**, 533–536, 1986.
110. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes - The Art of Scientific Computing*, Cambridge University Press; Cambridge, 2007.
111. R. E. Kalman, J. Basic Engineering **82**, 35–45, 1960.
112. S. Haykin, *Kalman Filtering and Neural Networks*; Wiley; New York, 2001.
113. T. B. Blank and S. D. Brown, J. Chemometrics **8**, 391–407, 1994.

A Molecular Dynamics Study of Peptide Stability in Mixed Solutions

Beate Moeser and Dominik Horinek

Institut für Physikalische und Theoretische Chemie, Universität Regensburg
93040 Regensburg, Germany

E-mail: dominik.horinek@chemie.uni-regensburg.de

The stability of biopolymers like proteins is decisively influenced by the solvent. In aqueous solutions, the presence of cosolutes is biologically exploited for the control of protein stability by stabilizing osmolytes or destabilizing denaturants. For the understanding of the molecular mechanism of the thermodynamic action of the cosolute on the protein stability, transfer models have been devised, in which the transfer free energies of folded and unfolded state between water and the cosolute solution are constructed as a sum of group contributions in a bottom-up approach. We present and discuss a simulation approach that is capable of evaluating the performance of different ways in which transfer models have been implemented.

1 Introduction: Protein Stability in Solutions

1.1 The experimental situation

Proteins in aqueous solutions can assume different conformational states. The native state is the state, which under natural conditions of temperature, pressure, and chemical environment is the thermodynamically most stable one[1]. When the conditions are altered thus that the native state is not stable anymore, a distinct structural transition occurs: the protein denatures from the compact, highly ordered folded state to the less organized denatured state. In the simplest case, in which no other states are populated, the equilibrium can be formulated as a thermodynamic two-state problem, in which the stability of the native, folded state in water is described by the unfolding Gibbs free energy

$$\Delta G_S^{N\to D} = G_S^D - G_S^N, \tag{1}$$

which is the difference of the free energies of the unfolded and the folded state. The index symbolizes the solvent environment in which the equilibrium takes places. We first look at a pure solvent S, for proteins this is typically water. When a cosolute is present as a third component, the folding equilibrium is affected, and thus a different unfolding free energy $\Delta G^{N\to D}$ is observed:

$$\Delta G_{CS}^{N\to D} = G_{CS}^D - G_{CS}^N, \tag{2}$$

where the index CS now symbolizes that we consider a mixture of solvent S and cosolute C. It is experimentally well established that $\Delta G_{CS}^{N\to D}$ depends linearly on the cosolute concentration c_C for many proteins[2]:

$$\Delta G_{CS}^{N\to D}(c_C) = \Delta G_S^{N\to D} + m \cdot c_C, \tag{3}$$

This is the defining equation for the "m value", which is a specific number for a given protein and a given cosolute[3]. Positive m values are characteristic for stabilizing cosolutes.

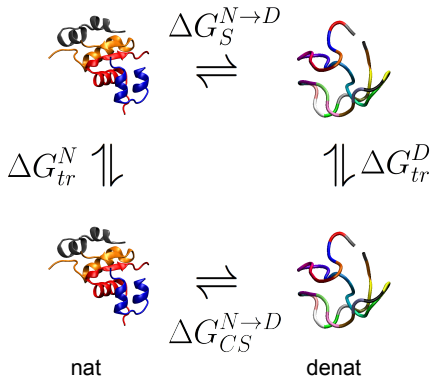


Figure 1: Transfer cycle of a protein in its native state (nat) and its denatured state (denat) in pure water solvent (S) and an cosolute/water mixture (CS).

A well known example is trimethylamine N-oxide, TMAO[4]. Negative m values lead to a destabilization of the folded state. Such cosolutes are known as denaturants, the probably best-known denaturant is urea[5].

1.2 Transfer models

The question we want to explore in the following is: can we explain the m value given that we know the sequence and the conformations of the folded and unfolded state of the protein? In a first step, a thermodynamic "transfer cycle" is constructed, which is shown in Figure 1. The thermodynamic cycle consists of two folding/unfolding processes in two different solvent environments and of two solvent transfers of the native and denatured state[6]. The sum of all four contributions along the transfer cycle must be zero. Therefore, the difference of two unfolding free energies can be expressed as the difference of two transfer free energies

$$m = \frac{\Delta G_{CS}^{N\to D}(c_C) - \Delta G_S^{N\to D}}{c_C} = \frac{\Delta G_{tr}^D - \Delta G_{tr}^N}{c_C}. \tag{4}$$

The transfer free energies (TFEs) ΔG_{tr} describe how the free energy of the protein in a given state changes during a hypothetical exchange of the solvent environment from pure water to a water/cosolute mixture. When this hypothetical exchange takes place between two solvents of equal volume, the resulting molarity-scale transfer free energies are related to changes in solvent-protein interactions[7,8]. The transfer processes are easier to deal with than the folding or unfolding, because their determination only requires the study of the protein in the folded and unfolded states independently. The task is then the determination of a transfer free energy of a given amino acid sequence in one given state. This is still not a simple task because of the molecular complexity of a protein and because of the manifold of possible conformations in protein structures. The interactions of the solvent and cosolute molecules with all protein groups are responsible for the stabilization or

destabilization of a given state. The question is now: can these contributions to the stability be disentangled, can we understand which groups of the protein interact favorably and which groups interact unfavorably with the cosolute? Unfortunately, there is no simple way how to disentangle all these contributions for a given protein conformation. The protein conformations have a complex three-dimensional shape, and it is unclear how strong the interactions with different groups in the protein are. One attempt to tackle this question is the introduction of the heuristic transfer model[9–11], in which additivity of the contributions of all amino acids in the protein is assumed.

$$\Delta G_{tr,protein} = \sum_X \Delta G_{tr,X}. \tag{5}$$

Furthermore, it is assumed that the backbone and side chain groups of every residue contribute independently to the transfer free energy (models, in which a finer atom-by-atom separation is made[12] will not be further discussed in this manuscript). The second assumption is that only groups that are exposed to the solvent contribute to the transfer free energy. Groups that are partially exposed contribute to the transfer free energy with a fraction relative to their solvent accessible area (ASA):

$$\Delta G_{tr,X} = \frac{ASA_X^{sc}}{ASA_X^{sc,ref}} \Delta G_{tr,X}^{sc,ref} + \frac{ASA_X^{bb}}{ASA_X^{bb,ref}} \Delta G_{tr,X}^{bb,ref}. \tag{6}$$

In this equation, reference values for the transfer free energy and the ASA of every group are introduced. The reference transfer free energies are scaled by the ratio of the solvent accessible area in the considered conformation and the reference solvent accessible area that describes the maximal exposure of this group. It is a conceptually simple step to combine the reference values in each of the two terms, which means simply to rewrite this equation as

$$\Delta G_{tr,X} = ASA_X^{sc} \left(\frac{\Delta G_{tr}}{ASA} \right)_X^{sc} + ASA_X^{bb} \left(\frac{\Delta G_{tr}}{ASA} \right)_X^{bb}. \tag{7}$$

This equation makes clear how the transfer model works: the reference values for each group define a ratio of transfer free energy per solvent accessible area (TFE/ASA), which gives one characteristic value for every group. Glycine contributes one backbone value to the transfer model, all other amino acids contribute two values to the transfer model, one for the backbone and one for the side chains. With 20 natural amino acids this results in a set of 39 TFE/ASA values, and it is important to realize how these numbers are determined in experiments. In this manuscript we do not go into much detail about the involved measurements and assume that transfer free energies for small molecules are available. In reality, their precise measurement is a very difficult experimental problem. Most experimental numbers that have been used are either approximations or are not very accurateIf we assume that the transfer free energies of monomeric amino acids are known, the TFE of the side chain of amino acid X is just the difference of the TFE of this amino acid and glycine:

$$\Delta G_{tr,X}^{sc} = \Delta G_{tr,X} - \Delta G_{tr,Gly}. \tag{8}$$

The determination of the TFE of the backbone is much more problematic. The TFE of glycine cannot be used, because it is heavily influenced by the end groups, even if a capped

amnio acid is used. Looking at sequences glycine, diglycine, triglycine etc. and determining incremental TFEs has the disadvantage that even these short peptides are rather flexible, and their conformation is not known. Thus, a relation of these TFEs to an ASA is unclear. These flexibility issues are avoided when cyclic diglycine is studied[13]. This molecule is strongly strained and has no significant conformational flexibility. Due to its cyclic nature it also has no end groups. Based on cyclic Gly_2, the transfer free energy of the backbone is determined by

$$\Delta G_{tr}^{bb} = \frac{\Delta G_{tr,cGly_2}}{2}. \tag{9}$$

Despite the use of this experimental number in applications of the transfer model, cyclic diglycine is not an optimal model compound, because the conformation of the peptide groups is rather unnatural: in proteins, the NH-CO groups are almost exclusively in a *trans* conformation, but in cyclic diglycine, they are in a *cis* conformation. We will come back to this problem later. The reference solvent accessible areas needed in the transfer model are obtained from structural informations of a large set of proteins in the protein data bank. The backbone and side chain ASAs are determined by locating Gly-X-Gly sequences in protein structures, isolating these tripeptide conformations from the protein structures, and determining their solvent accessible areas by the method of a rolling ball. After all the reference states have been determined, the transfer model can be applied to the unfolding of a given protein. The solvent accessible areas of all groups in the native state and in the denatured state are required. The conformations in the native state are known from X-ray or NMR experiments, and the group ASAs can be determined quite reliably. The structure of the denatured state is not as precisely known, and it is necessary to make assumptions about its structure. An often used choice is the model of Schellmann[14], which is an intermediate between a polymer in a good and in a poor solvent.

At this point it is important to note that the classical established transfer model[10] uses the same TFE for all backbone units, the one determined for glycine, However, the reference ASAs are different for all amino acids, because the backbone ASA largely varies with the side of the attached side chain. As a consequence, the characteristic TFE per ASA values of the backbone groups can differ by more than a factor of two between backbone groups belonging to different amino acid types. The different TFE/ASA values for the amino acid backbones are shown in Figure 2. All amino acids with side chains have a much larger TFE/ASA than glycine, which is the amino acid whose TFE is used in the derivation of all the TFE/ASA values. This simply reflects the fact that the backbone has a reduced solvent accessibility when a side chain is attached to it. The question is: should the reference TFE of the backbone be affected by the side chain next to it, which shields it partially from the solvent? This should be the case, if transfer free energies are related to solvent accessibilities. We extend the assumption of scaling of the TFE with the ASA to all backbone groups and require that the TFE/ASA should be equal for every backbone groups, regardless what the side chain attached to it is. This change defines the *universal backbone* transfer model, in which all backbone TFE/ASAs are treated consistently[15]. The group contributions are given by

$$\Delta G_{tr,X} = ASA_X^{sc} \left(\frac{\Delta G_{tr}}{ASA}\right)_X^{sc} + ASA_X^{bb}. \left(\frac{\Delta G_{tr}}{ASA}\right)^{bb}. \tag{10}$$

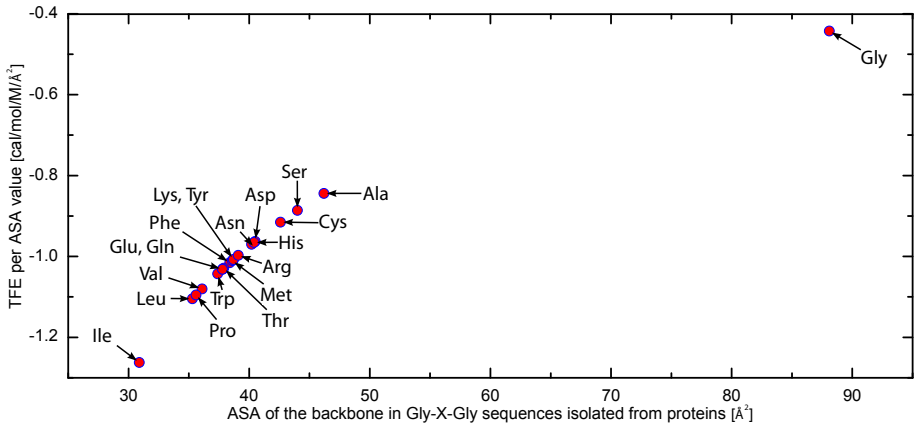


Figure 2: The reference state solvent accessible areas of the backbone group in the different amino acids and the resulting TFE/ASA values. The used reference state transfer free energy of the backbone is -193 /mol in one molar urea solution. It is evident that the resulting backbone TFE/ASA of glycine is very different from the backbone TFE/ASA of all other amino acids.

Note that the index 'X' in the backbone TFE/ASA term is now absent. This is only a sublet change in the defining equations, but now the model has only 21 reference values, one for every side chain, and one for all backbone groups. It is evident that the backbone contribution to the total transfer free energy in the universal backbone TM is much weaker than it is in the established model.

1.3 What can be learnt from simulations?

The predictive power of the different implementations of the transfer model requires two points. First, the underlying physical structure of this heuristic model must be a reasonable approximation to the actual thermodynamics. Second, good predictitions will only be possible if the reference values are appropriately chosen. Experimentally, these two points cannot be addressed independently, because the only results that are directly comparable to experiments are the total m value predictions for different proteins. Good predictions will be obtained if all requirements are fulfilled, but the possibility that different contributions of the model are only poor approximations cannot be recognized, if there are error cancellation effects and the final results for the m values is reasonable.

Molecular simulations offer a distinct chance for an in-depth test of different approximations made in the transfer model: states of a polypeptide chain that would not be accessible in experiments can be studied, and also peptide sequences that are not accessible experimentally can be studied. The strategy of our simulations is to compare simulation results for different peptide chains or for the different conformations of the same peptide chain with predictions based on the transfer model. Since all numbers are determined within the same simulation setup and force field, the quality of the predictions directly correlates to the quality of the transfer model formalism. Because there is no direct comparison made to experimental data, the quantitative agreement of the used force field num-

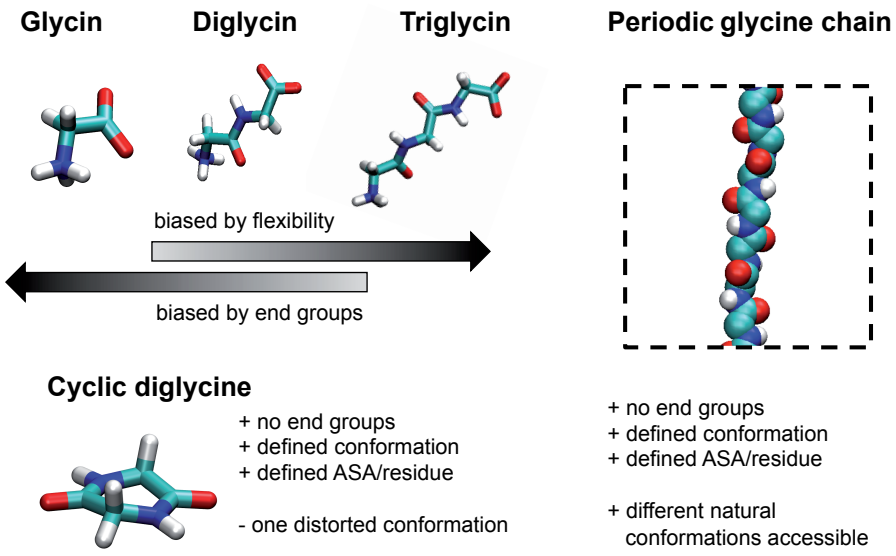


Figure 3: Model compounds for the determination of the reference transfer free energy of the backbone with their advantages and disadvantages.

bers and experimental numbers, which is usually a serious concern, is not necessary. The peptide models systems we pick are periodic peptide chains[16]. At first glance, they seem to be rather artificial model systems, but they combine several distinct advantages. They have no end groups because of their period nature. Therefore, only the residues present in the sequence contribute to the transfer free energies. The chains can be studied in states, in which all residues appear in similar conformations. In a single stretched peptide chain, all residues are maximally exposed to the solvent. Therefore, a stretched polyglycine chain is an optimal model system for the backbone. A summary of the different possible backbone models is shown in Figure 3. In a helix or in a β sheet, all residues are also in the same conformation and contribute the same to the transfer free energy. Thus the group contribution of every residue can be rigorously derived from the total transfer free energy.

2 Simulation Approach to Transfer Free Energies

There are various methods by which free energies of solvation of amino acids and peptides can be determined in combination with molecular simulation techniques[17–20]. Our goal is not to determine the solvation free energy itself, but the change in free energy upon change of the solvent composition. This is achieved by an approach that treats the polymer as an interface that is present in the solution. We now have to determine the change in free energy of this interface when the composition of the solvent is modified. The starting point is the Gibbs adsorption equation[21] for a one-dimensional interface of length L (later in this

chapter we will discuss also the case of a planar two-dimensional interface):

$$S^\sigma dT + Ld\tau + \sum_i N_i^\sigma d\mu_i = 0. \tag{11}$$

The excess particle numbers of component i, N_i^σ, are defined as[22]

$$N_i^\sigma = N_i - Vc_i, \tag{12}$$

where N_i and c_i are the total particle numbers of species i. The system volume is V, and the c_i is the bulk concentration of a solvent component i in the region of the simulation box that is unperturbed by presence of the interface. We restrict our discussion to isothermal processes, for which the first term in the Gibbs adsorption equation is zero. Division of the equation by L then yields for the interfacial tension change $d\tau$

$$d\tau = -\sum_i \Gamma_i d\mu_i, \tag{13}$$

where we have introduced the line excess at the polymer $\Gamma_i = N_i^\sigma/L$. We see that the change in Gibbs free energy of the interface induced by a change of the chemical potential of one solvent component is determined by the line excess Γ_i of this component times $d\mu_i$, the change of the components chemical potential. The line excesses Γ_i are calculated from an integral over the radial particle density profile $\rho_i(r)$

$$\Gamma_i = \int_0^\infty 2\pi r \left[\rho_i(r) - \rho_i^{\text{bulk}}\right] dr, \tag{14}$$

With the definition above, the polymer has no volume, as the total system volume is assigned to the bulk solvent phase. The line excess is in general different from zero even when a pure one-component solvent considered. The free energy changes that are determined are thermodynamically correct, but for the interpretation it is preferable to have a polymer with a finite volume and line excess of zero for a one-component solvent. This is achieved by defining a radius R of the polymer, thus giving it a volume. The radius R is determined by the condition that the pure solvent has no line excess at the polymer:

$$\Gamma_S(\text{pure solvent}) = 0 \tag{15}$$

with

$$\Gamma_i = \int_0^R 2\pi r \rho_i(r) dr + \int_R^\infty 2\pi r \left[\rho_i(r) - \rho_i^{\text{bulk}}\right] dr. \tag{16}$$

In our calculations, we use the cylindrical peptide-water and peptide-cosolute radial distribution functions g_{iP} instead of the radial densities ρ_i. They are normalized to one in the bulk solution,

$$\Gamma_i' = \int_0^R 2\pi r g_{iP}(r) dr + \int_R^\infty 2\pi r \left[g_{iP}(r) - 1\right] dr, \tag{17}$$

and the thus defined normalized surface excesses Γ_i' are simply given by the relation $\Gamma_i = c_i\Gamma_i'$. As we will show later, the change in line tension $d\tau$ is independent of the choice of R, and thus we are free to choose R according to the above condition. The preceding equations imply that an excess of any component has its origin in preferential attraction to or repulsion from the polymer chain. The line excesses vary in general with

the composition of the solution. For dilute solutions, the solvent contribution $\Gamma_S d\mu_S$ is often negligible, but we are interested here in cosolute concentrations up to several mol/l. Therefore, we cannot neglected the solvent term in our discussion. The transfer free energies we are interested have to be obtained as integrals over $d\tau$. The initial and final state are so far defined by the chemical potentials μ. The initial state should be the pure solvent, which implies

$$\mu_S = \mu_S^*, \tag{18}$$

and all other chemical potentials vanish. For a solution of N components, there are N chemical potentials that contribute to $d\tau$. We discuss here the simplest case, a ternary system with polymer, solvent and one cosolute species. In this case the chemical potential changes of solvent (S) and cosolute (C) are related. We realize that in equilibrium the chemical potentials are equal throughout the system. We pick a virtual subsystem, the 'bulk region' of the solution, where physical properties like the individual concentrations c_i are not influenced by the presence of the polymer. In this bulk region, a binary system is present, for which the Gibbs-Duhem equation at constant p and T states

$$n_S d\mu_S = -n_C d\mu_C \tag{19}$$

or

$$d\mu_S = -\frac{c_C}{c_S} d\mu_C \tag{20}$$

with the bulk concentrations c_S and c_C. We obtain[23]

$$d\tau = -\left[-\Gamma_S \frac{c_C}{c_S} + \Gamma_C\right] d\mu_C. \tag{21}$$

In principle, this equation can be directly used for the calculation of the transfer free energy by a one-dimensional integration over $d\mu_C$. The disadvantage is that chemical potentials are inconvenient to use in experiments as well as in simulations. Techniques like grand-canonical simulations are necessary to ensure constant chemical potentials. It is much simpler to perform simulations with constant particle densities. We transform the integration over $d\mu_C$ into an integration over the cosolute concentration dc_C in molarity scale. This is achieved by expressing the chemical potential of the cosolute in the molarity scale

$$\mu_C = \mu_C^\infty + RT \ln c_C y_C \tag{22}$$

which yields

$$d\mu_C = \left(\frac{\partial \mu_C}{\partial c_C}\right)_{p,T} dc_C \tag{23}$$

$$= RT\left[\frac{1}{c_C} + \left(\frac{\partial \ln y_C}{\partial c_C}\right)_{p,T}\right] dc_C. \tag{24}$$

Insertion of the expression for $d\mu_C$ into the Gibbs adsorption equation yields

$$\frac{d\tau}{dc_C} = -RT\left[-\frac{\Gamma_S}{c_S} + \frac{\Gamma_C}{c_C}\right]\left[1 + \left(\frac{\partial \ln y_C}{\partial \ln c_C}\right)_{p,T}\right]. \tag{25}$$

This is the desired equation in which the change in free energy $d\tau$ is related to the change of the cosolute concentration dc_C[16]. At this point is is also easy to see that the choice of R has no influence on $d\tau$: when R is changed by ΔR, the change in Γ_S is $c_S \Delta R$ and the change in Γ_C is $c_C \Delta R$. These two terms cancel exactly in equation 25.

2.1 Planar interfaces: β sheets

In our study we also look at infinite beta sheets. There, the theory has to be slightly adjusted: instead of a line excess (particles per unit length) we have to determine the surface excess (particles per unit area). This is easily achieved by modifying equation 16 to

$$\Gamma_i = 2 \left\{ \int_0^R \rho_i(z) dz + \int_R^\infty \left[\rho_i(z) - \rho_i^{\mathrm{bulk}} \right] dz \right\}. \tag{26}$$

The integration variable is now the distance z from the interface, which is in the xy plane. The factor $2\pi r$ which reflected the cylindrical symmetry is replaced by a factor of 2, which reflects now that both sides of the β sheet are exposed to the solvent.

3 Simulation Methods

In the following we describe how simulations are used for the calculation of the individual terms in Eq. 25.

3.1 Simulation details

The KBFF force fields for urea[24], is combined with the SPC water model[25] for the simulations of binary mixtures of urea and water. The peptide chains are described by the GROMOS 53a6 force field[26]. All simulations are done with the Gromacs simulations program[27,28]. The long-range electrostatics are treated by the smooth particle-mesh Ewald summation with a real-space cutoff of 1.2 nm. A smooth van der Waals force cutoff is applied between 0.9 and 1.0 nm. Simulation times vary between 50 ns and 150 ns for every system. The system temperature is 300 K, the pressure is 1 bar. For the bulk simulations, the pressure is maintained by isotropic pressure coupling, for the peptide simulations the z-component is excluded from the pressure coupling, because it controls the conformation of the peptide chain. For the simulation of β sheets, the x and y component are fixed.

3.2 Line excesses of solvent and cosolute

For the determination of the line excesses Γ_S and Γ_C, simulations of ternary system solvent/cosolute/peptide are performed. The peptide chain is oriented along the z-axis of the system. The simulations are performed at constant temperature and pressure. Pressure scaling is only performed in the x and y direction, because the z-length of the box is fixed because it keeps the polymer at the desired length per monomer. The bulk concentrations of solvent and cosolute are not equal to the total number of particles divided by the system volume. They have to be determined in a region far from the polymer. We determine the

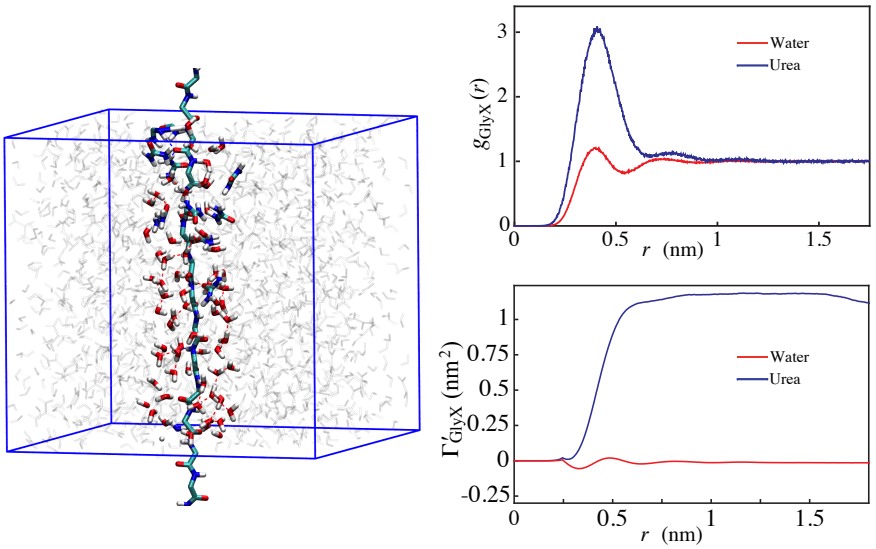


Figure 4: Left: Snapshot of an extended polyglycine chain in a urea/water mixture. For clarity, only the first solvation shell of the peptide s shown, the other solvent molecules are shown as shades. Right: radial distribution functions and running integrals of the line excesses of urea and water around a stretched polyglycine chain. The bulk urea density is $c_S = 4.3$ mol/l.

cylindrical radial densities of both solution species. Because of the cylindrical geometry of the polymer conformation, nicely shaped functions are obtained that decay to a bulk value within a few solvation shells. In the shown example in Figure 4 bulk conditions are reached for $r > 1$ nm. For larger values a constant bulk density is reached, which we extract as the desired bulk concentrations c_C and c_S. Normalization of the two densities to a bulk value of one yields the (cylindrical) radial distribution functions g_{SP} and g_{CP}.

3.3 The bulk enhancement factor

The terms discussed so far are interfacial properties of the system, which have their origin in how the solvent and cosolute behave in the polymer's vicinity. The thermodynamic description is completed by a term that describes the bulk mixing thermodynamics of the binary solvent/cosolute mixture (see equation 25)

$$\left(\frac{\partial \ln y_C}{\partial \ln c_C}\right)_{p,T}. \tag{27}$$

This quantity is accessible from simulations of the binary system. A convenient approach is the application of Kirkwood-Buff theory (KBT)[29,30], which relates thermodynamic properties of the system to structural properties. The central quantities of this theory are the Kirkwood-Buff integrals G_{ij}, which are integrals over the radial distribution functions[24,31]

$$G_{ij} = \int_0^\infty 4\pi r^2 \left[g_{ij}(r) - 1\right] dr. \tag{28}$$

The Kirkwood-Buff expression for the desired derivative of the activity coefficient is

$$\left(\frac{\partial \ln y_C}{\partial \ln c_C}\right)_{p,T} = \left(\frac{1}{1 + c_C(G_{CC} - G_{CS})} - 1\right). \tag{29}$$

There are a few points that need to be considered when KBT is used in combination with molecular dynamics simulation results. First, Kirkwood-Buff theory is defined for systems that are coupled to particle baths and in which the particle numbers can fluctuate. The grand canonical radial distribution functions differ in a subtle, but distinct way from the canonical radial distribution functions. Second, equation 28 is only valid for $R \to \infty$, which cannot be achieved with simulation data. Without going into further detail, we note that these problems have been addressed by several authors[32–34] and that the method can be applied with reasonable accuracy. In our case, the bulk enhancement is calculated from 23 simulations of 60 ns length each, in which the urea concentration is varied between 0 and 11.32 mol/l. The results are fitted to an analytical function, which yields for the SPC water model and the KBFF urea force field at 1 bar and 300 K:

$$c_C \cdot \left(\frac{d \ln y_C}{d \ln c_C}\right)_{p,T} = 0.00017572 c_C + 0.058176\sqrt{c_C} - 0.15394. \tag{30}$$

whre c_C is measured in mol/l.

4 Results

In the following, we present results from simulations of peptides in urea water mixtures and their implications for the transfer model. The transfer model introduces severe assumptions about the composition of the transfer free energy of a protein. In recent years there have been several studies that demonstrate the limitations of group additivity and ASA scaling[35–37].

4.1 Group Additivity

Our intention here is the study of additivity in a very special context: what are the errors when peptide transfer free energies are approximated as the sum the transfer free energies of the individual amino acid residues? Of course, peptide sequences and conformations are too numerous to be studied completely. As a starting point, we pick two examples and study if additivity is satisfied.

4.1.1 Glycine and Tyrosine

We first test the combination of glycine with the hydrophobic residue tyrosine. The simulations of polyglycine and polytyrosine give transfer free energies per residue of -513 and -836 J/mol per M urea. Periodic chains with the sequence $(\mathrm{GlyTyr})_4$ and $(\mathrm{Gly_2Tyr_2})_2$ are studied. The additivity prediction for both sequences is a TFE per residue of -677 J/mol per M urea. The simulations result, -689 and -678 J/mol per M urea are in very good agreement, showing that the additivity is well satisfied in this case.

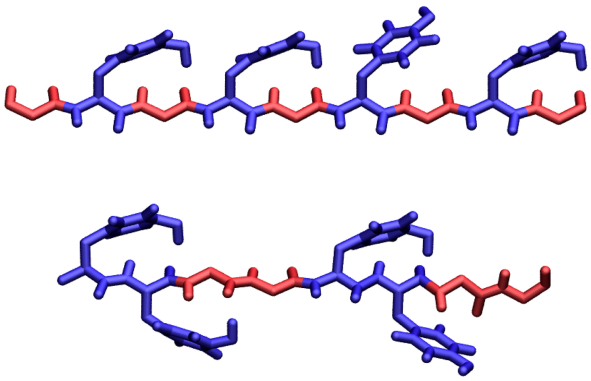


Figure 5: The mixed peptides $(\mathrm{GlyTyr})_n$ and $(\mathrm{Gly_2Tyr_2})_n$ used for testing the additivity of transfer free energies.

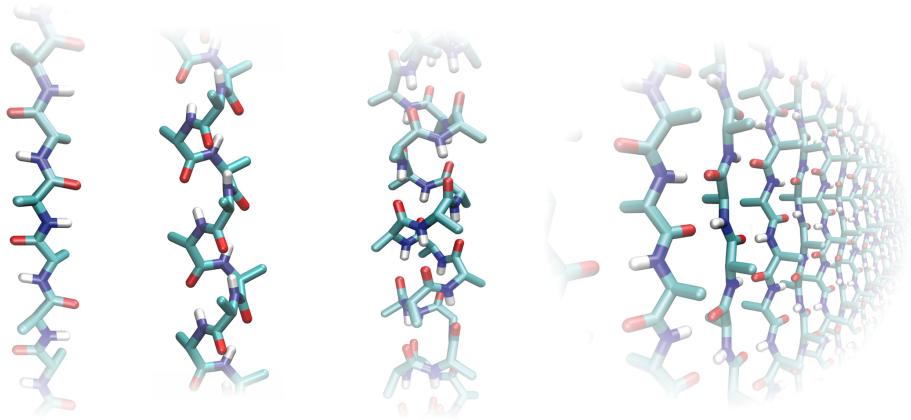


Figure 6: Four folding motifs of a polypeptide chain used for the evaluation of the TFE/ASA scaling hypothesis: extended chain, 3_{10} helix, α helix, β sheet.

4.1.2 Glycine and Serin

For glycine and tyrosine, only the sequence $(\mathrm{Gly_2Ser_2})_n$ is studied. The additivity prediction of -469 J/mol per M urea underestimates the simulation result of -530 J/mol per M urea. The relative error of 12% is significant, but still tolerable.

5 Transfer Free Energies of Secondary Structure Motifs - ASA Scaling

In this section we address the second assumption of the TM, the dependence of the group TFEs on their solvent accessible area. While this concept is applicable to any type of conformation, we restrict ourselves to periodic secondary structures, in which all residues give

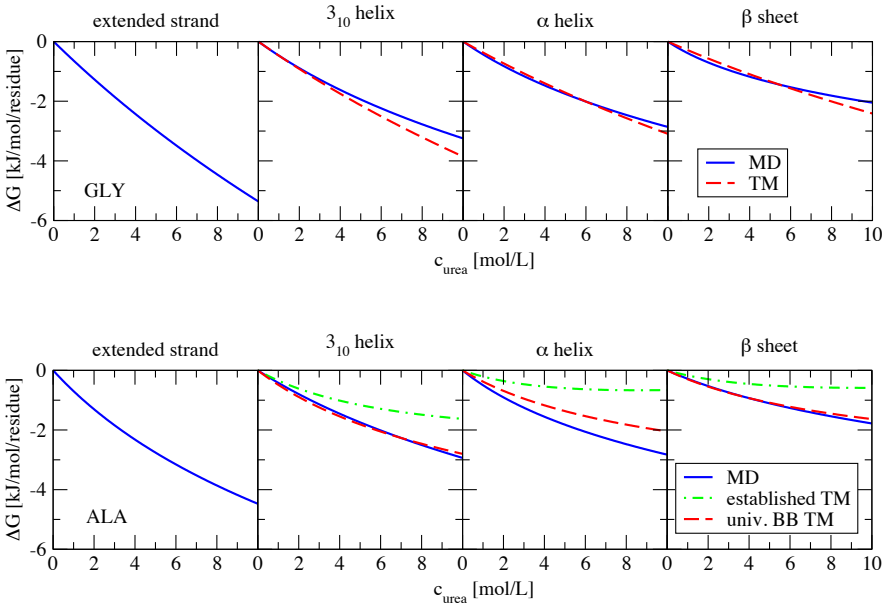


Figure 7: Transfer free energies of polyglycine and polyalanine in different conformations determined by explicit MD simulations and predicted by the transfer model based on the results of the extended strands as reference state.

identical contributions, and the assignment of the total TFE onto the different residues is transparent. Four different structural motifs are studied: extended chain, 3_{10} helix, α helix, and a β sheet, as shown in Figure 6. First we consider polyglycine chains. Glycine has no side chain, and therefore only the backbone groups enter. Since the conformations are chosen such that all residues are in identical conformations, the transfer model reduces to

$$\Delta G_{\mathrm{Gly}} = \frac{\Delta G_{\mathrm{Gly}_N}}{N} = ASA^{bb}_{Gly} \left(\frac{\Delta G_{tr}}{ASA} \right)^{bb}_{Gly}. \tag{31}$$

For the validation of the ASA scaling we now determine the TFE and the ASA in two states 'A' and 'B' and determine the prediction of $\Delta G(B)$ according to

$$\Delta G(B) = \frac{ASA^{bb}_{Gly}(B)}{ASA^{bb}_{Gly}(A)} \Delta G(A). \tag{32}$$

The extended chain serves as the reference state 'A' in our simulations. Fig. 7 shows the simulated TFEs as a function of urea concentration for the three 'folded' structures and their TM-prediction based on the extended chain TFE, which is also shown in the left panel. The simulations show that ASA scaling works well for the polyglycine secondary structures. The mean relative error is 12%. We now turn to amino acids with side chains.

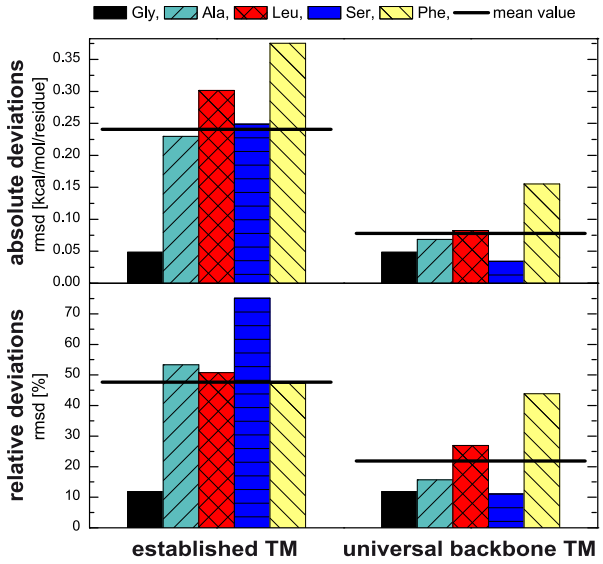


Figure 8: Absolute and relative errors of the predictions based on the established transfer model (left) and the universal backbone transfer model (right) for the transfer free energies of five different homopeptides in different secondary structures.

For these chains we have to consider backbone and side chain contributions

$$\frac{\Delta G_{X_N}}{N} = ASA_X^{sc}\left(\frac{\Delta G_{tr}}{ASA}\right)_X^{sc} + ASA_X^{bb}\left(\frac{\Delta G_{tr}}{ASA}\right)_X^{bb} \tag{33}$$

The backbone contributions are determined from the backbone groups ASAs and from the simulation results of the polyglycine chain. According to the classical implementation of the transfer model, the backbone contribution is derived as

$$\Delta G_X^{bb}(B) = \frac{ASA_X^{bb}(B)}{ASA_X^{bb}(A)}\Delta G_{Gly}^{bb}(A) \tag{34}$$

which yields the prediction for the folded peptide X based on the extended chain reference data of the amino acids X and Gly

$$\Delta G_{\mathrm{X}}(B) = \frac{ASA_X^{sc}(B)}{ASA_X^{sc}(A)}\Delta G_X^{sc}(A) + \frac{ASA_X^{bb}(B)}{ASA_X^{bb}(A)}\Delta G_{Gly}^{bb}(A) \tag{35}$$

The predictions of the folded TFEs of the four amino acids alanine, leucine, serin, and phenylalanine that are obtained with Equation 35 give a significant error. The average relative error obtained for this four amino acids and glycine is shown in Figure 8, each in the three different folds, is approximately 40%, which means that even semi-quantitative predictions cannot be achieved. We now seek the root of the bad performance. In principle, the heuristic model might simply not work. However, it is remarkable that the glycine predictions are reasonable, and only the predictions of the other studied amino acids are poor. We put our focus on the treatment of the backbone and study the universal backbone

transfer model, which replaces the backbone ASA of the stretched amino acid X in the denominator, $ASA^{bb}_X(A)$, by the backbone ASA of stretched glycine $ASA^{bb}_{Gly}(A)$:

$$\Delta G^{bb}_X(B) = \frac{ASA^{bb}_X(B)}{ASA^{bb}_{Gly}(A)} \Delta G^{bb}_{Gly}(A). \tag{36}$$

The prediction for the transfer free energy reads now

$$\Delta G_{\mathrm{X}}(B) = \frac{ASA^{sc}_X(B)}{ASA^{sc}_X(A)} \Delta G^{sc}_X(A) + \frac{ASA^{bb}_X(B)}{ASA^{bb}_{Gly}(A)} \Delta G^{bb}_{Gly}(A) \tag{37}$$

Figure 8 shows the predictions that are obtained with the universal backbone transfer model next to the results of the established transfer model. It is immediately evident that the predictions of the universal backbone model are significantly better than the results obtained with the established model. The average absolute error is reduced by more than a factor of two, the relative error is reduced to about 20%. The accuracy is good enough that semi-quantitative predictions seem possible. The obvious interpretation of these simulation results is that the universal backbone TM is superior to the established implementation. It is not only a consistent extension of the underlying mathematical model, it also gives a dramatic improvement of the predictions of transfer free energies of different secondary structure motifs.

6 Conclusions

We present a molecular dynamics approach that is able to quantitatively evaluate relations between transfer free energies and the structure of polypeptide chains. Using this approach, we are able to assess transfer models that are based on additive group transfer free energies and on these group's solvent accessible areas. Our results show that the transfer model works reasonably, but only after a fully consistent treatment of the backbone contributions has been implemented. The previously established implementation is shown to give poor agreement. In the end we want to shortly mention the next step: applying the universal backbone TM as we have tested it so far to the cases for which transfer models were originally designed: the folding of real proteins. In fact, it was possible to show that the conceptually well working universal backbone TM is also giving good predictions of experimental m values of proteins based on the currently best collection of group reference data[15]. The interpretation arising from this analysis is that backbone and side chain contributions to the denaturation of proteins by urea are on average equal, which is in agreement with several other studies[38,39]. On the other hand, application of the previously established transfer model, which performs poorly in the evaluation by MD simulations, results in the prediction that denaturation by urea is driven almost exclusively by effects of the backbone[11,40]

References

1. CB Anfinsen, *Principles That Govern Folding of Protein Chains*, Science, **181**, 223–230, 1973.
2. CN Pace and KL Shaw, *Linear extrapolation method of analyzing solvent denaturation curves*, Proteins, **41**, 1–7, 2000.
3. CN Pace, *The Stability of Globular Proteins*, CRC Critical Reviews in Biochemistry, pp. 1–43, 1975.
4. PH Yancey and JF Siebenaller, *Trimethylamine oxide stabilizes teleost and mammalian lactate dehydrogenases against inactivation by hydrostatic pressure and trypsinolysis.*, J Exp Biol, **202**, 3597–3603, 1999.
5. PL Whitney and C Tanford, *Solubility of amino acids in aqueous urea solutions and its implications for the denaturation of proteins by urea.*, J Biol Chem, **237**, 1735–1737, 1962.
6. C Tanford, *Protein denaturation. C. Theoretical models for the mechanism of denaturation.*, Adv Prot Chem, **24**, 1–95, 1970.
7. A Ben-Naim, *Standard thermodynamics of transfer. Uses and misuses*, J Phys Chem, **82**, 792–803, 1978.
8. B Moeser and D Horinek, *The role of the concentration scale in the definition of transfer free energies*, Biophys Chem, **196**, 68–76, 2015.
9. C Tanford, *Isothermal Unfolding of Globular Proteins in Aqueous Urea Solutions*, J Am Chem Soc, **86**, 2050–2059, 1964.
10. M Auton and DW Bolen, *Predicting the energetics of osmolyte-induced protein folding/unfolding*, Proc Natl Acad Sci USA, **102**, 15065–15068, 2005.
11. M Auton, LMF Holthauzen, and DW Bolen, *Anatomy of energetic changes accompanying urea-induced protein denaturation*, Proc Natl Acad Sci USA, **104**, 15317-15322, 2007.
12. EJ Guinn, LM Pegram, MW Capp, MN Pollock, and MT Record, *Quantifying why urea is a protein denaturant, whereas glycine betaine is a protein stabilizer*, Proc Natl Acad Sci USA, **108**, 16932–16937, 2011.
13. M Auton and DW Bolen, *Additive transfer free energies of the peptide backbone unit that are independent of the model compound and the choice of concentration scale*, Biochemistry, **43**, 1329–1342, 2004.
14. JA Schellman, *Protein Stability in Mixed Solvents: A Balance of Contact Interaction and Excluded Volume*, Biophys J, **85**, 108–125, 2003.
15. B Moeser and D Horinek, *Unified Description of Urea Denaturation: Backbone and Side Chains Contribute Equally in the Transfer Model*, J Phys Chem B, **118**, 107-114, 2013.
16. D Horinek and RR Netz, *Can Simulations Quantitatively Predict Peptide Transfer Free Energies to Urea Solutions? Thermodynamic Concepts and Force Field Limitations*, J Phys Chem A, **115**, 6125–6136, 2011.
17. TP Straatsma and JA McCammon, *Computational Alchemy*, Annu Rev Phys Chem, **43**, 407–435, 1992.
18. P Kollman, *Free-Energy Calculations - Applications to Chemical and Biochemical Phenomena*, Chem Rev, **93**, 2395–2417, 1993.
19. H Kokubo, RC Harris, D Asthagiri, and B Montgomery Pettitt, *Solvation Free Energies of Alanine Peptides: The Effect of Flexibility*, J Phys Chem B, **117**, 16428–16435, 2013.

20. MR Shirts and VS Pande, *Solvation free energies of amino acid side chain analogs for common molecular mechanics water models*, J Chem Phys, **122**, 134508, 2005.
21. RJ Hunter, *Foundations of Colloid Science 2nd Ed*, Oxford Univ Pr, Oxford, 2000.
22. HJ Butt, K Graf, and M Kappl, *Physics and Chemistry of Interfaces*, Wiley-VCH, Berlin, 2003.
23. VA Parsegian, RP Rand, and DC Rau, *Osmotic stress, crowding, preferential hydration, and binding: a comparison of perspectives*, Proc Natl Acad Sci USA, **97**, 3987–3992, 2000.
24. S Weerasinghe and PE Smith, *A Kirkwood-Buff Derived Force Field for Mixtures of Urea and Water*, J Phys Chem B, **107**, 3891–3898, 2003.
25. HJC Berendsen, JPM Postma, WF van Gunsteren, and J Hermans in *Intermolecular Forces*, Pullmann, B., Ed.; Reidel, Dordrecht, 1981.
26. C Oostenbrink, A Villa, AE Mark, and WF van Gunsteren, *A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6*, J Comput Chem, **25**, 1656–1676, 2004.
27. E Lindahl, B Hess, and D van der Spoel, *GROMACS 3.0: a package for molecular simulation and trajectory analysis*, J Modeling, **7**, 306–317, 2001.
28. B Hess, C Kutzner, D van der Spoel, and E Lindahl, *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*, J Chem Theory Comput, **4**, 435–447, 2008.
29. JG Kirkwood and FP Buff, *The Statistical Mechanical Theory of Solutions. I*, J Chem Phys, **19**, 774–777, 1951.
30. A Ben Naim, *Inversion of the Kirkwood–Buff theory of solutions: application to the water–ethanol system*, J Chem Phys, **67**, 4884–4890, 1977.
31. PE Smith, *Chemical Potential Derivatives and Preferential Interaction Parameters in Biological Systems from Kirkwood-Buff Theory*, Biophys J, **91**, 849–856, 2006.
32. P Krüger, SK Schnell, D Bedeaux, S Kjelstrup, TJH Vlugt, and JM Simon, *Kirkwood–Buff Integrals for Finite Volumes*, J Phys Chem Lett, **4**, 235–238, 2013.
33. SK Schnell, X Liu, JM Simon, A Bardow, D Bedeaux, TJH Vlugt, and S Kjelstrup, *Calculating Thermodynamic Properties from Fluctuations at Small Scales*, J Phys Chem B, **115**, 10911–10918, 2011.
34. P Ganguly and NFA van der Vegt, *Convergence of Sampling Kirkwood–Buff Integrals of Aqueous Solutions with Molecular Dynamics Simulations*, J Chem Theory Comput, **9**, 1347–1355, 2013.
35. DS Tomar, V Weber, BM Pettitt, and D Asthagiri, *Conditional Solvation Thermodynamics of Isoleucine in Model Peptides and the Limitations of the Group-Transfer Model*, J Phys Chem B, **118**, 4080–4087, 2014.
36. T Hajari and NFA van der Vegt, *Peptide Backbone Effect on Hydration Free Energies of Amino Acid Side Chains*, J Phys Chem B, **118**, 13162–13168, 2014.
37. J Heyda and J Dzubiella, *Thermodynamic Description of Hofmeister Effects on the LCST of Thermosensitive Polymers*, J Phys Chem B, **118**, 10979—10988, 2014.
38. DR Canchi and AE Garcia, *Cosolvent Effects on Protein Stability*, Annu Rev Phys Chem, **64**, 273-–293, 2012.
39. DR Canchi and AE Garcia, *Backbone and Side-Chain Contributions in Protein Denaturation by Urea*, Biophys J, **100**, 1526–1533, 2011.
40. M Auton, J Rösgen, M Sinev, LMF Holthauzen, and DW Bolen, *Osmolyte effects on protein stability and solubility: A balancing act between backbone and side-chains*, Biophys Chem, **159**, 90–99, 2011.

Modelling and Analysis of Interfacial Structural Transitions in Ionic Liquids at Charged Interfaces

Vladislav V. Ivaništšev[1,2], Kathleen Kirchner[1], Tom Kirchner[1], Sean O'Connor[1], and Maxim V. Fedorov[1]

[1] Department of Physics, Scottish Universities Physics Alliance (SUPA), Strathclyde University, John Anderson Building, 107 Rottenrow East, Glasgow, G4 0NG UK.

E-mail: maxim.fedorov@strath.ac.uk

[2] Institute of Chemistry, University of Tartu, Ravila 14A, 50114 Tartu, Estonia

In these lecture notes we will briefly overview current progress in theory and simulations of (room temperature) **i**onic **l**iquids (**IL**s) at charged interfaces (CIs). ILs are important *highly concentrated* electrolyte media for many applications - from energy storage to friction. When a certain IL is to be chosen for a given application, a suitable one has to be selected from, practically, an infinite number of possible combinations of cations and anions. Therefore, the main role or modelling and theoretical techniques in this area is to understand the main *trends* in the IL behaviour at CIs and reveal the most important *guiding factors* that determine it. The notes are focussed on computational and theoretical analysis of the fundamental properties of the structure of the electrical double layer (EDL) in ILs that is dramatically different from the one in diluted electrolytes. Several methodological aspects that are specific to IL systems are discussed.

1 Introduction

Interfacial effects in (room temperature) ILs at charged surfaces are very important for IL applications in electrochemistry[1,2], energy storage[3], catalysis[4] and other areas (such as e.g. lubrication[5,6]). However despite of the many works published on this subject (several hundreds only in 2014) there is still no general agreement in the literature about the main factors that govern the structure and properties of ILs at charged interfaces. This is mainly due to a large amount of available combinations of ionic liquids (and their mixtures!) and different surfaces[1].

In these notes we will briefly overview current progress in theory and simulations of ILs at CIs with a focus on methods of description and analysis of interfacial structural transitions in ILs. We will present a critical (qualitative) analysis of available molecular modelling and experimental data using recently developed concept of surface charge compensation that allows one to compare results for different ILs on the same methodological footing.[7] Several methodological aspects that are specific to IL systems will be also discussed.

The notes are largely based on our recent works on modelling interfacial structural transitions in ILs[7–10].

2 Motivation

ILs are important *highly concentrated* electrolyte media for many applications - from energy storage to friction[1]. When a certain IL is to be chosen for a given application, a

suitable one has to be selected from, practically, an infinite number of possible combinations of cations and anions[11,12,1]. Therefore, the main role or modelling and theoretical techniques in this area is to understand the main *trends* in the IL behaviour at CIs and reveal the most important *guiding factors* that determine it.

Charge/voltage driven structural transitions in the EDL in ionic liquids (ILs) have recently attracted large interest in experimental[13,5,14–17], theoretical[18] and computational[19–21,7,22,23,8–10] communities due to the importance of this subject for a variety of IL applications[24,1]. Such, it has been shown recently that the ion layering/packing determines the free energy barriers for a solute approaching the electrode[25–27], ion self-diffusion in charged (slit-)nanopores[28], and the potential-dependent variation of nanoscale-level friction at the IL–electrode interfaces[5,29] (see also the recent Editorial in the J. Phys. Chem. C by Kornyshev and Qiao[30]).

Bazant et al.[18] suggested that general trends in structural transitions in ILs upon surface charging are determined by the crossover between the over-screening and the crowding regime in the EDL structure. In a recent modelling work[7] it was shown that this crossover corresponds to a structural transition from a multilayer (intermediate charges) to an over-crowded structure (high charges; superposition of two or more counter-ion layers) through the formation of a monolayer structure at a certain charge density value. In the Refs[8–10] it has been suggested that these trends may be universal and are expected to be found in many IL systems.

In these notes we will show that molecular modelling techniques combined with proper theoretical analysis can make the next step towards rationalising general mechanisms of charge-driven interfacial structural transitions in ILs. To illustrate the main ideas behind the methods we will present several case studies that investigate and compare interfacial structural behaviour of different coarse-grained[31,32,7] and fully atomistic[9,26] models of ILs.

3 Interfacial Structures in Ionic Liquids at Charged Surfaces

It has been shown in a number of experimental, modelling and theoretical works that ILs undergo structural reorganisation on molecular level upon charging the interface. This has direct effects on mass and charge transfer happening at the interface[24] as well as on the differential capacitance (C_{diff}) dependence on the electrode potential[32,31,18] and temperature[33,34]. The interfacial structure of ionic liquids also influences lubrication properties of the interface.[29,35] For recent reviews on experimental studies in this area we refer to Refs.[36–38]; for an overview of molecular modelling studies see Ref.[22].

From a practical point of view, the response of the interfacial ionic liquid structure to changes of the surface charge can be important for a variety of applications of ionic liquids such as supercapacitors, batteries and electrocatalysis[4,2,3]. Many independent experimental studies that used different techniques (high-energy X-ray reflectivity, nuclear magnetic resonance, surface apparatus, and atomic force microscopy (AFM)) reported formation of multilayered structures of ionic liquids at charged interfaces with alternate layers of cations and anions (see e.g. Refs[39–42]). Similar multilayered interfacial structures in ionic liquids were also found in a number of Molecular Dynamics (MD) studies that used different simulation methods (force fields and simulation conditions), ionic liquid models (coarse-grained and atomistic) as well as electrode models (carbon material, structureless charged surface, metal, etc)[22].

It was shown that charging of the surface affects the magnitude of the ion layering at the interface[32,31,25,14,29,35]. Such, recent AFM studies have shown that the molecular-level structure and, consequently, interfacial properties of ionic liquids at charged surfaces can be changed by varying the surface potential[29,35]. It has been also shown in these works that the interfacial structure can be changed by varying chemical composition of the ionic liquids, i.e. varying the length of alkyl chains of organic cations or anion type, etc.

However, we note that a multilayer-like structure is not the only possible arrangement of ions at the interface. Indeed, monolayer-like interfacial structures of ionic liquids at charged surfaces has been reported for several ionic liquids systems[43]. Critical analysis of several studies applying the sum frequency generation spectroscopy technique led the author to a conclusion that ions at the solid–liquid interface in ionic liquids are organized into essentially one ionic layer – a structure that resembles the classical Helmholtz picture of an ionic monolayer at a charged surface[44], where the structural correlations between the ions beyond this layer resemble the bulk liquid structure[43].

Trying to resolve the apparent contradiction between the different views on the interfacial structure of ionic liquids at charged surfaces recently Kirchner *et al.* performed a MD study of structural reorganization in ionic liquids upon surface charging using a wide scale of surface charges[7]. The authors of Ref.[7] used a coarse-grained model of ionic liquids that was previously developed in Refs.[31,32]. The results of this work have shown that varying the surface charge one can observe both multilayer structure and monolayer structure in ionic liquids. Although the conclusions were derived from a rather crude model that describes ionic liquid cations and anions as charged Lennard-Jones spheres we think that they can be applied when analysing interfacial properties of more complex models and real systems.

Figure 1 schematically presents two main types of the molecular-scale structure of the electrified electrode–ionic liquid interface that can be formed in a model ionic liquid – the multilayer structure (LEFT) and the monolayer structure (RIGHT). The simulation setup for the MD simulations and the model parameters are described in details in Refs[31,32,7] (see also below). The model systems consist of two smooth charged surfaces (11 nm$\times$ 11 nm) and a variable number of ionic pairs (up to 1050)[31,32,7,45]. Simulations were carried out using molecular dynamics at temperatures from 250 to 500 K, in an NVT ensemble. The results shown in Figures 1 and 2 are drawn for the symmetric ionic liquid model with a bulk number density of 0.32 ions per nm^3 and ionic radii of 0.5 nm.

The multilayer structure on Figure 1:LEFT is characterized by segregation of ions into cationic and anionic layers; that type of interfacial structure is related to the overscreening phenomena in ionic liquids[31,32,46,18]. The overscreening implies that the innermost layer contains more charge than is needed to compensate the surface charge (σ). The excess of charge in the innermost layer (λ_1) is compensated by the following ionic liquid layers in the transition region. The interfacial layering of ionic liquid particles is reflected in the oscillation of the electrostatic potential (ϕ) (Figure 1), ionic number density (Figure 2A), ionic charge density and its derivatives, like the excess of charge in the i-th layer (λ_i, Figure 2C). Thus, in the multilayer regime, the molecular-scale structure of ionic liquid at a charged interface can be divided into three regions: the *innermost layer* of counter-ions in direct contact with the surface, the *transition zone* of several nanometer thickness that consists of alternating ion layers and the *bulk-like* structured ionic liquid at larger distances from the surface.

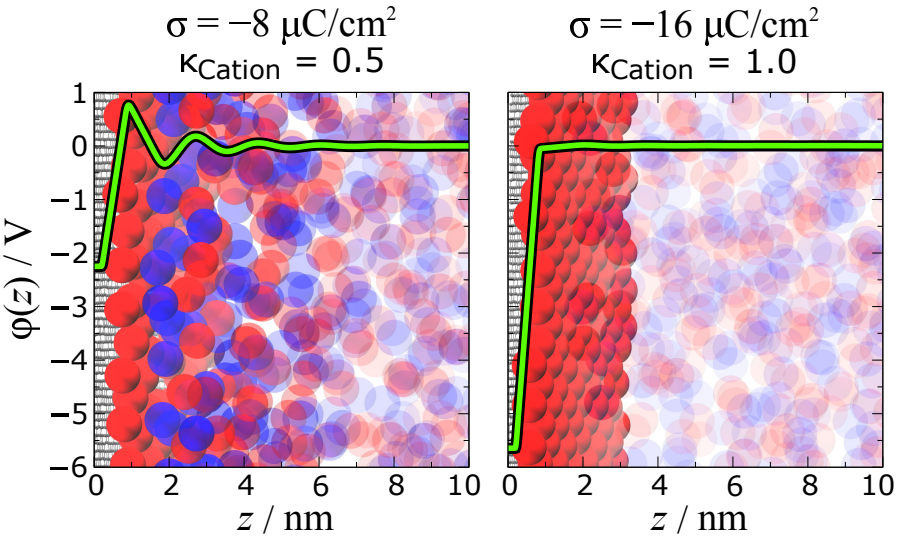


Figure 1: Electrostatic potential $\phi(z)$ across two *model* interfaces representing *multilayer* (LEFT) and *monolayer* (RIGHT) structures in a model ionic liquid. For the sake of simplicity, here both anions and cations are represented by spherical particles of the same size. In the underlaying simulation snapshots the electrode, cations and anions are depicted as gray, red and blue spheres, respectively. The opacity level of the snapshots is adjusted proportionally to the variation of the amplitude of $\phi(z)$. The figure is taken from Ref.[8]. Copyright 2014, The Electrochemical Society (ECS).

We note that, contrary, there is no transition zone in the case of the monolayer structure. Consequently, there are no oscillations in the electrostatic potential (Figure 1:RIGHT) – the innermost layer coexists directly with the bulk-like structured ionic liquid.

4 Analysis and Description of Interfacial Structural Transitions in Ionic Liquids

To perform a quantitative analysis of structural transitions in the interfacial region we introduce a parameter $\theta^{\max}$ that corresponds to the maximum charge density that can be stored in a monolayer for a selected type of ionic liquid ions (that would correspond to a densely packed monolayer of ions at physically possible maximum of ion packing density)[7]. By its definition, for a given ionic liquid $\theta^{\max}$ depends only on the ion charge and geometric parameters of the selected ion type (size and shape).

The essential features of the monolayer structure are – (i) a linear electrostatic potential drop across the interface; (ii) the charge density stored in the monolayer equals to $\theta^{\max}$; (iii) the electrode counter-ions that form the monolayer completely compensate the surface charge, i.e. $\sigma = \theta^{\max}$. Due to the fact that $\theta^{\max}$ depends only on the charge and geometry of the ions, one can use the surface charge renormalized by $\theta^{\max}$ to compare different ionic liquid systems on the same universal footing. The dimensionless surface charge compensation parameter (κ) is then defined as:

$$\kappa_{\mathrm{Ion}} = \left| \frac{\sigma}{\theta_{\mathrm{Ion}}^{\max}} \right| . \tag{1}$$

Figure 2 illustrates the relationships between the structure of the transition region and the packing of the innermost layer of counter-ions for different electrode charge densities. Figure 2A shows the dependency of the ion number density (ρ_N) on the distance from the electrode (z) and the renormalized surface charge density – in the form of a contour map of $\rho_N(z, \kappa)$. As can be seen from this figure, vertical valleys divide the interface into distinct regions: the innermost layer, the transition region and the bulk ionic liquid.

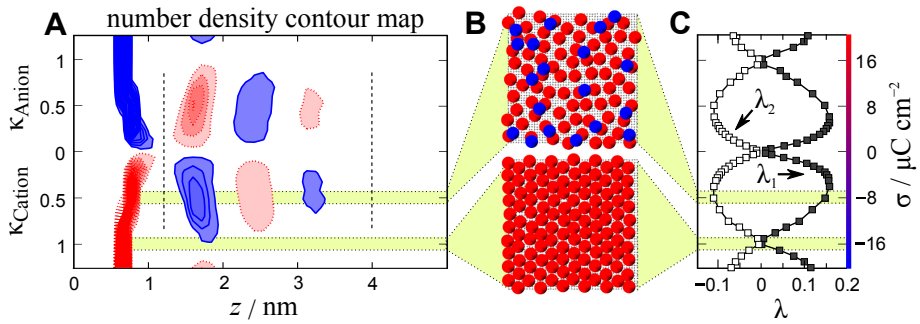


Figure 2: **A. LEFT:** Dependency of anion (dark, blue) and cation (light, red) number density (ρ_N) on the distance from the electrode (z) and renormalized surface charge density (κ_{Ion}). Dashed lines divide the interface into the innermost, transition and bulk regions. The contour interval equals to ρ_{bulk}, the first contour starts at $1.5\rho_{\mathrm{bulk}}$, and the $\rho_N(z, \kappa_{\mathrm{Ion}})$ peaks are cut at $7.5\rho_{\mathrm{bulk}}$ to facilitate the visual analysis. **B. MIDDLE:** The TOP subfigure shows dislocation of anions (blue) and cations (red) within 1.7 nm from the surface (gray) at $\sigma = -8\ \mu\mathrm{C\ cm}^{-2}$ and $\kappa_{\mathrm{Cation}} = 0.5$. The BOTTOM subfigure shows dislocation of ions within 1.7 nm from the surface (gray) at $\sigma = -16\ \mu\mathrm{C\ cm}^{-2}$ and $\kappa_{\mathrm{Cation}} = 1.0$. **C. RIGHT:** Excess of charge in the first (λ_1, filled marks) and in the second (λ_2, empty marks) interfacial layers. The κ_{Ion} values of 1 correspond to $\pm 16\ \mu\mathrm{C\ cm}^{-2}$ on the σ-scale shown on the left. Areas corresponding to the *multilayer* and *monolayer* structures are highlighted in color and with dotted lines. The figure is taken from Ref.[8]. Copyright 2014, The Electrochemical Society (ECS).

The number of layers in the transition region (within 1–4 nm for the coarse-grained model) depends on the κ_{Ion}. It increases with increasing the surface charge until $\kappa_{\mathrm{Ion}} \approx 0.5$ and then decreases revealing a bare monolayer at $\kappa_{\mathrm{Ion}} = 1$. The horizontal valleys at $\kappa_{\mathrm{Ion}} = 1$ confirm the coexistence of the monolayer with bulk-like ionic liquid.

Although the interfacial structure can exhibit quasi-crystal ordering in the vicinity of zero surface charge,[47] in the presented simulations the interchange of cations and anions happens disorderly upon crossing the zero charge mark ($\kappa = 0$). Onwards, the innermost layer becomes pressed against the surface by the strong electrostatic attraction between counter-ions and the charged surface. This is evident from the thinning of the first layer and monotonous grow of the $\rho_N(z, \kappa)$ of counter-ion with increasing κ (2A).

For comparison, in Figure 2B there are shown anions (blue) and cations (red) within 1.7 nm from the charged surface at $\kappa_{\mathrm{Cation}} = 0.5$ (TOP) and $\kappa_{\mathrm{Cation}} = 1$ (BOTTOM). As can be seen from this figure, already at $\kappa_{\mathrm{Cation}} = 0.5$ the innermost layer is dominated by the electrode counter-ions. In previous works observations of similar structures were

associated with the *lattice saturation* effect and the maximum of differential capacitance curve (C_{diff} vs. E).[48] However, the figure shows that within the distance of 1.7 nm from the electrode there are also co-ions coordinated with several (2–4) counter-ions. Thus, we think that the correlations between cations and anions actually define the distance between the counter-ions in the innermost layer at $\kappa_{\mathrm{Ion}} = 0.5$.

The packing density of ions at the interface can be further increased upon surface charging without crowding up to the monolayer structure formation (from $\kappa_{\mathrm{Ion}} = 0.5$ to $\kappa_{\mathrm{Ion}} = 1.0$). To our best knowledge the corresponding regime of electrostriction was not previously described in the literature, although similar ideas have been discussed in terms of C_{diff} vs. $|\Delta\phi|^{-n}$ dependence[49–51,46].

The transition between the multilayer structure to the monolayer structure is illustrated using the charge excess in the i-th layer (λ_i) in Figure 2C. The parameters λ_i is defined through the cumulative charge density as follows:

$$\lambda_i = \kappa_{\mathrm{Ion}} \times \left(\left| \frac{cn_Q(z_i)}{\sigma} \right| - 1 \right), \tag{2}$$

where z_i corresponds to an extremum (or a step height) of $|cn_Q(z)/\sigma|$ on the interval that defines the i-th layer. $cn_Q(z) = \frac{1}{A}\int_0^z \rho_Q(z')\mathrm{d}z'$, where $\rho_Q(z) = \rho_N^{\mathrm{Cation}}(z) \cdot q_{\mathrm{Cation}} + \rho_N^{\mathrm{Anion}}(z) \cdot q_{\mathrm{Anion}}$. Note that the electrostatic potential quantities $\rho_Q(z)$, $cn_Q(z)$ and λ are derived from ionic number densities.

The increase of the absolute charge excess in the interfacial layers with increasing κ_{Ion} from 0 to 0.5 manifests the transition from the disordered to the multilayer structure, while the decrease of the charge excess indicates the vanishing of the multilayer structure towards exposure of the monolayer structure at $\kappa_{\mathrm{Ion}} = 1$. The λ_i vs. κ_{Ion} dependence represents an analog of dimensionless "reaction coordinate" for the reorganization process.

The screening factor ($|cn_Q(z)/\sigma|$) as well as similar conceptions of charge magnitude and the normalized surface charge density were previously employed in several previous works[32,20,52,53].

On the larger κ-scale the monolayer structures in the coarse-grained models are formed at integer values of κ_{Ion} (adjusted for compressibility),[45] similarly as for one-dimensional lattice Coulomb gas model (with lower fugacity).[54] An exact solution for the latter model also indicates that the maximal charge layering happens at renormalized surface charge densities of $\sim 0.3 \div 0.5$. The presented results reveal that the similar is true for three-dimensional ionic liquids models.

The recent work of Ivaništšev et al[9] has shown that the multilayer to monolayer structural transitions observed in simulations by coarse-grained IL models[7,8,10] can be still found in fully atomistic simulations of ILs at CIs. In the recent atomistic MD simulations of Paek et al. and Hu et al. one can also see evidences of the monolayer formation in the case of complex atomistic ionic liquid models (see Figure 4 in Ref.[55] and Table II in Ref.[20]). Therefore, the κ-scale can be used for rationalising general trends in more realistic systems. Figure 3 shows the relationship between ionic size and maximum packing charge density (formed at $\kappa_{\mathrm{Ion}} = 1$). As a first approximation, it is assumed here that the ions are roundish and non-polarizable (we note that the former assumption is true for the alkali, halogenide and, perhaps, PF_6^- and BF_4^- ions). The presented $\theta^{\max}$ values are approximate and can be adjusted by considering compressibility and specific packing of ions (that should be important for larger ions, which in addition are electrochemically

more stable due to presence of stabilizing functional groups and larger distance from the electrode[56,57]).

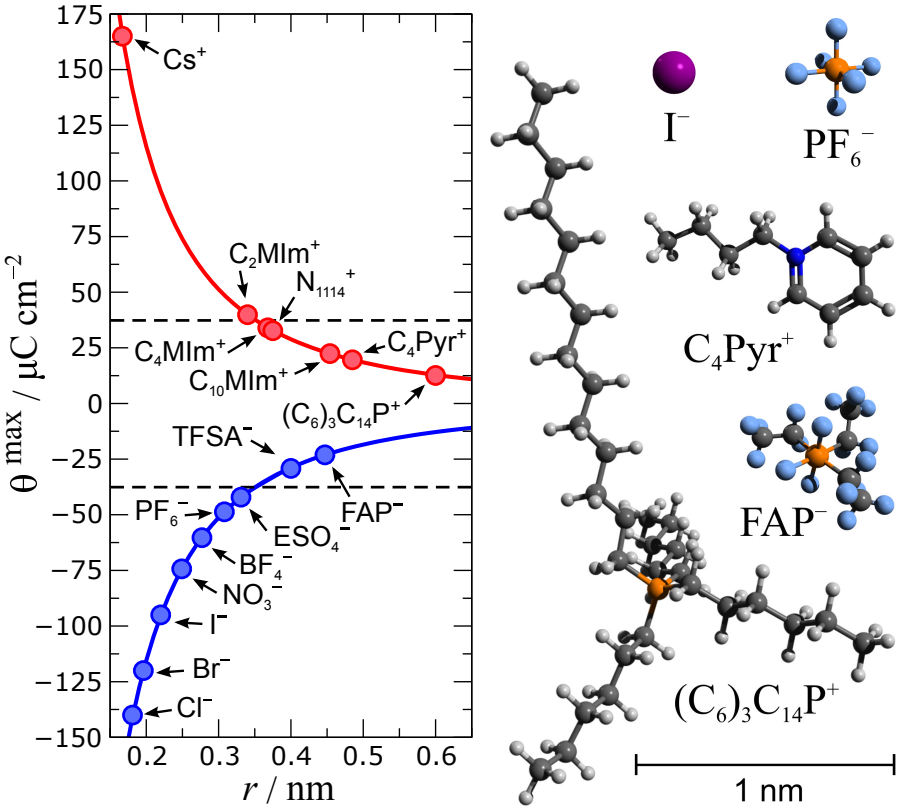


Figure 3: LEFT: Dependence of sterically determined maximum charge density ($\theta_{\mathrm{Ion}}^{\mathrm{max}}$) on ionic radius ($r$). Dashed lines denote limiting values of surface charge achievable in electrochemical experiments with ionic liquids. For complex ions the radii were estimated from molecular volumes taken from Refs.[58,59] assuming that ions are spherical in shape. The radii of halogenide anions and alkali metals cations were taken from Ref.[60]. RIGHT: ball-and-stick models of chosen ions amplified relative to the presented scale. The figure is taken from Ref.[8]. Copyright 2014, The Electrochemical Society (ECS).

Complicated geometric shape of ions, specific interionic interactions and (partial) charge transfer between the ions and the charged surface are present at real interfaces. According to Figure 3 the formation of monolayer structure can be achieved for a number of relatively large ions: $\mathrm{C}_n\mathrm{MIm}^+$ and $\mathrm{C}_n\mathrm{MPyr}^+$ with $n \geq 4$, TFSA^-, FAP^-, etc, for which the listed effects can be neglected in a first approximation.

5 Molecular-Scale Structural Reorganization of Ionic Liquids in Response to the Surface Charge

In this section we present a unified view on charge-driven structural transitions in the electrical double layer in ionic liquids and summarise molecular-scale mechanisms of the ionic liquid structural response to the surface charge.

Bazant et al.[18] suggested that general trends in structural transitions in ILs upon surface charging are determined by the crossover between the over-screening and the crowding regime in the EDL structure. In the modelling works discussed above[7,8] it was shown that this crossover corresponds to a structural transition from a multilayer (intermediate charges) to an overcrowded structure (high charges; superposition of two or more counter-ion layers) through the formation of a monolayer structure at a certain charge density value. Below we show that the conclusions from the Refs.[18,7,8] are not effects of an oversimplified view on ILs and can be used for describing realistic systems. We illustrate that by comparing structural behaviour of three different coarse-grained IL models[31,32,7] with the behaviour of a fully atomistic model of 1-butyl-3-methylimidazolium tetrafluoroborate ([BMIm][BF_4])[26]. The section is largely based on our recent paper on this subject[9].

5.1 Simulations

All simulations were performed using the classical Molecular Dynamics (MD) method in the NVT ensemble at a temperature of 350 K with the Gromacs 4.5.5 software[61]. The simulation setups (equilibration, length of simulations, system parameters and computational methods) were overall the same to the ones used in our previous works: Ref.[26] (fully atomistic model of [BMIm][BF_4]) and Refs.[7] (coarse-grained models). Periodic boundary conditions were applied in all directions. The cut-off of the Lennard-Jones interactions was taken to be 2.6 nm with shifted potential method. The long-range Coulomb interactions were handled by the Particle-Mesh Ewald (PME) method[62] with a cutoff of 2.9 nm, a grid spacing of 0.112 nm and corrected for slab geometry with using the Yeh and Berkowitz correction method for using 3D Ewald methods in slab-like geometries, 3D-EWC.[63]. The neighbour list for non-bonded interactions was updated for every 10^{th} integration step.

5.1.1 Coarse-grained simulations

The simulated systems represent model ionic liquid (IL) ions confined between two model electrodes[31,32,7]. The models represented IL ions as charged Lennard-Jones spheres[31,7] with the cation-to-anion diameter (d_{LJ}) ratios were chosen to be $1:1$ (large anion – LA), $1:0.8$ (medium anion – MA) and $1:0.5$ (small anion – SA) with $d_{\mathrm{LJ}}(\mathrm{Cation}) = 1.0\,\mathrm{nm}$ (the SA model was previously used in][31,7]). The electrodes consist of 2500 Lennard-Jones spheres with a diameter (d_{LJ}) of 0.22 nm that are arranged on a square lattice with a size of $11\,\mathrm{nm} \times 11\,\mathrm{nm}$, in x and y directions. Distance between the electrodes was chosen to be 54 nm, 36 nm and 24 nm for the LA, MA and SA systems, respectively. The electrodes were charged with surface charge densities (σ) ranging from 0 to $\pm 24\,\mu\mathrm{C/cm^2}$ for the LA model, from 0 to $\pm 40\,\mu\mathrm{C/cm^2}$ for the MA model, and from 0 to $\pm 76\,\mu\mathrm{C/cm^2}$ for the SA model. The ion pair number was fixed in all simulations to be equal 1050. We used the leap-frog algorithm for integrating Newton's equations of motion with a 0.01 ps time step.

Velocity rescaling was used with a temperature coupling constant of 1.0 ps[64]. The force field parameters of the models are summarized in Table 1.

Table 1: Force field parameters of the coarse-grained models of ionic liquids and the electrodes. Ions and electrode particles are modelled as charged Lennard-Jones spheres with only repulsive part of the short-range potential $U_{\mathrm{LJ}}^{ij} = 4\varepsilon_{\mathrm{LJ}}\left[(d_i + d_j)/r\right]^{12}$, where $\varepsilon_{\mathrm{LJ}} = 1.8698\ \mathrm{kj\ mol^{-1}}$. The arithmetic combination rule was used for both d_{LJ} and $\varepsilon_{\mathrm{LJ}}$. Similar to the previous works[31,32,7] the relative permittivity was set to 2 to account for polarization effects.

Particle	d_{LJ} / nm	Interaction	d_{LJ} / nm
Cation	1.00	Large anion–Cation	1.00
Large anion (LA)	1.00	Medium anion–Cation	0.89
Medium anion (MA)	0.79	Small anion–Cation	0.75
Small anion (SA)	0.50	Cation–Electrode	0.60
Electrode	0.22	Large anion–Electrode	0.60
		Medium anion–Electrode	0.49
		Small anion–Electrode	0.35

5.1.2 Fully atomistic simulations

The system consisted of two rigid graphene slabs with dimensions of 3.408 nm by 3.4433 nm separated by a distance of 10.4 nm. 374 [BMIm][BF_4] ion pairs were placed between these surfaces and equilibrated. The OPLS-AA force field was used together with partial charges taken from[65] for the IL. The charges were screened by a factor of 0.79 to account for electronic polarisability[26].

5.1.3 Iterative systems preparation

Preparation and analysis of the simulations for all IL models were done using the NaRIBaS toolbox [a]. To improve the simulation convergence and statistical sampling, five independent molecular configurations (replicas) per system were generated with the Packmol software[66]. For every combination of IL model and the surface charge, the results then were averaged over productive simulation runs for all five replicas of the system. The system preparation and simulation consisted of the following steps, repeated iteratively for all IL models, surface charges and replicas[67]:

1. Energy minimization
   (integrator steep, 10000 steps, electrodes neutral);
2. NVT simulation with a small time step
   (integrator md, $dt = 2$ fs, 100000 steps = 200 ps, electrodes neutral);
3. Energy minimization
   (integrator steep, 10000 steps, electrodes charged);
4. NVT simulation with a small time step
   (integrator md, $dt = 2$ fs, 100000 steps = 200 ps, electrodes charged);

[a] K. Kirchner and M.V. Fedorov, NaRIBaS: A scripting framework for computational modeling of Nanomaterials and Room temperature Ionic liquids in Bulk and Slab; Version 1.0 (2013). http://sourceforge.net/projects/naribas/

5. NVT simulation with a large time step
   (integrator md, $dt = 10\,\mathrm{fs}$, 3500000 steps = 35 ns, electrodes charged).

5.2 Analysis

As in the previous section, we used the unified κ-scale. Below we use the κ-scale for generalised analysis with a focus on the regions $0 < \kappa_{\mathrm{Anion}} < 1$ and $0 < \kappa_{\mathrm{Cation}} < 2$, where κ_{Cation} corresponds to the negative surface charge density values ($\sigma < 0$) and κ_{Anion} corresponds to the positive values ($\sigma > 0$).

Because the monolayer structure is characterised by smearing of oscillations in the electrostatic potential $\phi(z)$-profiles[7,8], the value of ($\theta_{\mathrm{Ion}}^{\mathrm{max}}$) was extracted from the simulation results at the point of surface charge that corresponds to a linear potential drop. The potential drop at $\kappa_{\mathrm{Ion}} = 1$ can be roughly approximated as:

$$\phi_{\mathrm{ML}} = \frac{d}{\epsilon}\theta_{\mathrm{Ion}}^{\mathrm{max}}, \tag{3}$$

where $\theta_{\mathrm{Ion}}^{\mathrm{max}} \approx e\frac{q_{\mathrm{Ion}}}{r_{\mathrm{Ion}}^2}$, d is the distance between the surface and the monolayer charge planes, r_{Ion} is ionic radius, q_{Ion} is ionic charge, e is elementary charge, and ϵ is permittivity of the monolayer structure. $\theta_{\mathrm{Cation}}^{\mathrm{max}}$ was found to be the same for all three coarse grain-systems ($+16\,\mu\mathrm{C/cm}^2$) in accordance with the fact that the cation model is the same in all systems. This value equals to the density of one cation per $1\,\mathrm{nm}^2$ of the surface that corresponds to the dense coverage of the surface by the cations. $\theta_{\mathrm{Anion}}^{\mathrm{max}}$ values were found to be -68, -26 and $-16\,\mu\mathrm{C/cm}^2$ for the systems with small, medium and large anions, respectively. These values also correspond to the dense coverage of the surface by the anions. For the atomistic model of [BMIm][BF_4], $\theta_{\mathrm{Anion}}^{\mathrm{max}}$ is $-100\,\mu\mathrm{C/cm}^2$, $\theta_{\mathrm{Cation}}^{\mathrm{max}}$ is $+38\,\mu\mathrm{C/cm}^2$.

The restructuring process at different charge densities can be illustrated with the use of the parameter (λ_i) defined above. In the analysis below we will consider only the λ_i parameter of the first ion layer (λ_1), therefore the index i is omitted.

As was discussed in the previous section, the κ_{Ion}–scale represents a universal analogue of dimensionless "reaction coordinate" for the EDL restructuring process in response to the surface charge. Analysis of the dependence of λ on κ_{Ion} allows to study the evolution of the EDL structure in terms of the charge excess. Namely, an increase of the charge excess in the first interfacial layer (λ_1) manifests formation of a multilayer EDL structure, while the decrease of the charge excess indicates the vanishing of the multilayer EDL structure towards the formation of the monolayer structure at $\kappa_{\mathrm{Ion}} = 1$.

5.3 Simulation results

The electrostatic potential $\phi(z)$-profiles were calculated using the solution of the Poisson equation:

$$\phi(z) = \phi(0) - \frac{\sigma z}{\epsilon\epsilon_0} - \frac{1}{\epsilon\epsilon_0}\int_0^z z'\rho_Q(z')\mathrm{d}z', \tag{4}$$

where $\rho_Q(z)$ is the charge density at distance z from the electrode surface, ϵ is the relative permittivity and ϵ_0 is the electric constant. Value of $\phi(0)$ was chosen so that the

electrostatic potential in bulk IL equals 0, and ϵ was set to 2 in order to account for the polarization effect. The potential drop across the modelled interfaces can be then calculated as $U = \phi(L/2) - \phi(0)$, where L is a distance between the model electrodes.

In Figures 4 and 5, $\phi(z)$-profiles for the LA model and an atomistic [BMIm][BF_4] model[26,9] are shown, respectively. The striking similarity in the profiles for these two different types of IL models indicates that the overscreening-to-crowding crossover[18] happens via a state of complete surface charge compensation by the counter-ions of the first layer[7]. Taking this state as a reference point, we define a unified κ-scale, on which σ is normalized by the maximum charge density stored in the first interfacial layer ($\theta_{\mathrm{Ion}}^{\mathrm{max}}$):

$$\kappa_{\mathrm{Ion}} = \left| \frac{\sigma}{\theta_{\mathrm{Ion}}^{\mathrm{max}}} \right| . \tag{5}$$

Regarding the results for the atomistic model, it should be clarified that oscillations of the electrostatic potential are seen at all surface charges, even for an uncharged electrode surface. Particularly, at $\kappa = 0$ and $\kappa = 1$ the oscillations reflect the heterogeneity of the bulk IL[68]. Moreover, two linear regions in the potential drop (at $\kappa_{\mathrm{Cation}} = 1.4$ in Fig. 5) are related to existence of two distinct orientations of imidazolium ring relative to the surface plane[20]. Despite these nuances, from the $\phi(z)$-profiles it is seen that for both type of models the oscillations in $\phi(z)$-profiles dissapear upon $\kappa \to 1$ and reappear at $\kappa > 1$. The

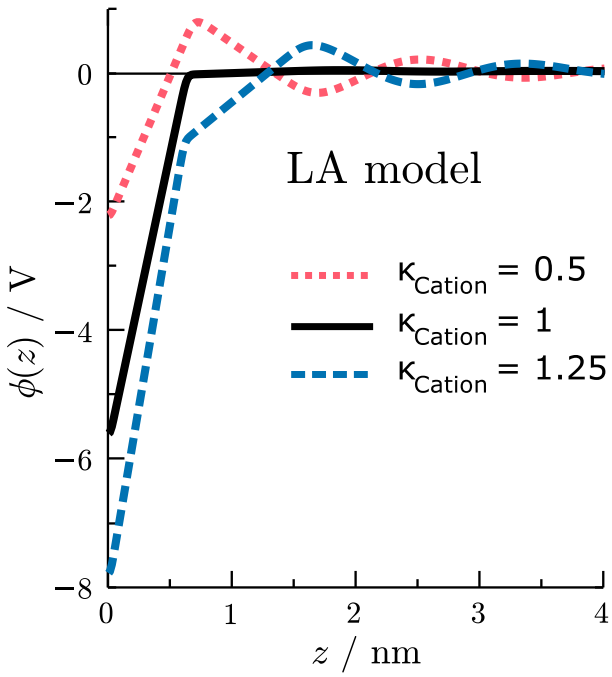


Figure 4: Electrostatic potential profiles ($\phi(z)$), in the model system with large anion (LA), indicate transition from the multilayer structure (in the so called overscreening regime, $\kappa < 1$) to the crowded structure (in the so called crowding regime, $\kappa > 1$) through formation of the monolayer structure at $\kappa = 1$. The figure is taken from the Supporting Materials to Ref.[10].

dumping is associated with the decrease in the interfacial ion layering magnitude. This is evident from the dependence of charge excess in interfacial layers (λ) on κ discussed below.

Figure 6 presents the dependency of the IL ion number density ρ_N from the distance to the electrode z and κ in the form of $\rho_N(z, \kappa)$ contour maps. These maps illustrate charge-dependent layering of cation (light, red) and anion (dark, blue) for the coarse-grained (left, MA) and the atomistic (right) model IL systems.

Although the compared IL models are quite different from each other, in both cases the contour maps reveal similar features of the IL structural response to the surface charge that are described below.

As can be seen, the vertical ridges of high ion number density divide the interfacial region into distinct regions of ion accumulation. We refer to the region of counter-ions accumulation closest to the electrode, as the *first* layer. Counter-ions $\rho_N(z, \kappa)$ in the first layer grows upon surface charging while counter-ions become pressed against the surface due to the strong electrostatic attraction. Differently, in the subsequent layer the $\rho_N(z, \kappa)$ grows until some saturation at $\kappa_{\mathrm{Ion}} \approx 0.5$ and then decreases until $\kappa_{\mathrm{Ion}} = 1.0$. The dotted horizontal lines point to the areas of practical absence of layering around $\kappa_{\mathrm{Ion}} = 1.0$ (Figure 6).

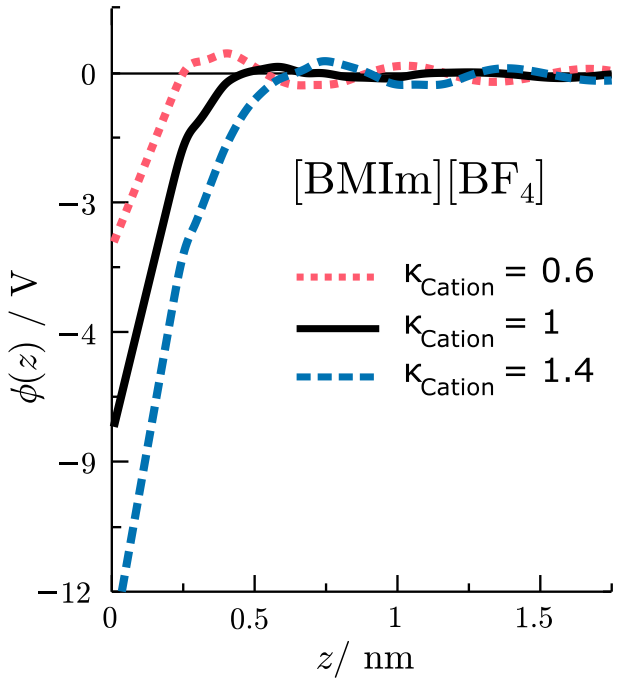


Figure 5: Electrostatic potential profiles ($\phi(z)$), at the model graphene–[BMIm][BF_4] interface, indicate the transition from the multilayer structure (in the so called overscreening regime, $\kappa < 1$) to the crowded structure (in the so called crowding regime, $\kappa > 1$) through formation of the compact-layer structure at $\kappa = 1$. The figure is taken from the Supporting Materials to Ref.[10].

To facilitate comparison of *different* IL systems, in Figure 7, we plot λ versus κ for cationic and anionic layers. As can be seen, the evolution of the EDL structure upon surface charging is qualitatively the same for all coarse-grained and atomistic models. This implies that the main mechanisms of ion accumulation at the surface are governed mostly by electrostatic interactions and sterical effects. Yet, despite of the apparent *general* similarity seen in Figures 6 and 7, there is a lamination of the $\rho_N(z, \mathrm{Cation})$ at large κ_{Cation} due to the presence of both parallel and perpendicular orientation of the $[\mathrm{BMIm}]^+$ ring in the first cationic layer.

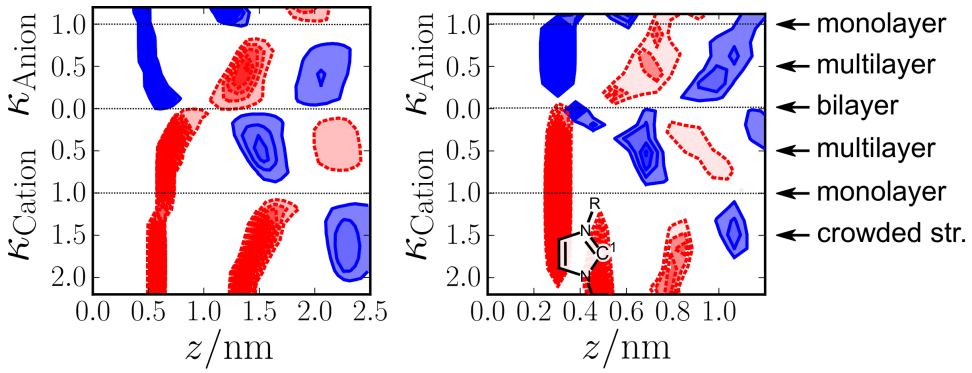


Figure 6: The figure presents the ion number density $\rho_N(z, \kappa)$ contour maps that illustrate charge-dependent layering of cation (light, red) and anion (dark, blue) for the MA (left) and the [BMIm][BF_4] (right) models. The contour interval equals to ρ_{bulk}, the first contour starts at $1.5\rho_{\mathrm{bulk}}$ (MA) and $2.5\rho_{\mathrm{bulk}}$ ([BMIm][BF_4]), and the $\rho_N(z, \kappa)$ peaks are cut at $7\rho_{\mathrm{bulk}}$ to facilitate the visual analysis. For [BMIm][BF_4] the positions of the IL anions and cations are assigned to the centres of C^1 and B atoms respectively; the picture of $[\mathrm{BMIm}]^+$ indicates that the lamination of the $\rho_N(z, \mathrm{Cation})$ at large κ_{Cation} values is due to the presence of parallel and perpendicular (shown) orientation of the aromatic ring in the first cationic layer at high surface charges. The figure is taken from Ref.[9]. Copyright 2014, Elseveir.

6 Poly(a)morphic Portrait of the Electrical Double Layer in Ionic Liquids

As a summary, Figure 8 illustrates general mechanisms of structural transitions in the EDL of ILs at flat interfaces which are represented as an ensemble of charge-dependent poly(a)morphic structures[9]. These mechanisms can be described by the formation and mutual transformation of *ionic bilayer* (D,E), *multilayer* (C,F) and *monolayer* (B,G) structures followed by *crowding* of the counter-ions at high surface charges (A,H).

- **D,E:** Ionic bilayers of cations and anions that are formed at small surface charges $\kappa \approx 0$. Analysis of the MD simulations results reveals that at small absolute σ values the first layer consists of two correlated subsystems – anionic and cationic – that form the ionic bilayer. Upon surface charging, the cationic and anionic subsystems become separated in space due to depletion of the co-ions and enrichment of the counter-ions near the surface. The anionic and cationic subsystems become completely separated by $\kappa \approx 0.5$ when the multilayer structure is formed.

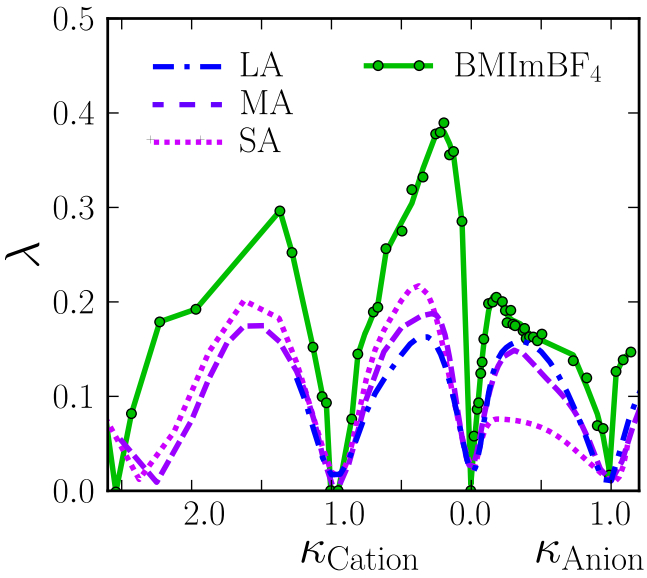


Figure 7: Variation of the normalised charge excess in the first interfacial layer (λ) with κ for the four model ILs studied in this work. The figure illustrates the overall similarity of the charge-induced EDL restructuring in these *different* model ILs. The κ_{Ion} values of 1.0 correspond to the $\theta_{\mathrm{Ion}}^{\mathrm{max}}$ values of the charge density σ. λ minimum at $\kappa \approx 2.3$ indicates the formation of a crowded layer of cations which accommodates more counter-ions that is expected from a superposition of two dense monolayers; that happens due to the squeezing and reorientation of the cations in the strong electric field at these high charges. The figure is taken from Ref.[9]. Copyright 2014, Elseveir.

- **C,F:** Multilayered structures formed at intermediate κ-values that are composed from well-distinguishable layers of cations and anions that alternate in the normal direction (see Figure 6). These multilayered structures are characterised by maximal values of the charge excess parameter (λ) (see Figure 7).

- **B,G:** Cationic (B) and anionic (G) monolayer structures formed at κ-values close to 1.0. In the cationic monolayer most of the $[\mathrm{BMIm}]^+$ rings lie parallel to the surface and few of the $[\mathrm{BMIm}]^+$ rings orient perpendicular to the surface. Noteworthy that the reorientation of a small part of the $[\mathrm{BMIm}]^+$ rings happens for the same reason as the displacement of the tails from the surface – in order to provide higher charge density in the ionic layer closer to the surface. The tails are pulled away from the surface at $\kappa > 0.5$ forming a flexible sublayer. This observation is in agreement with experimental evidences of formation of a similar intermediate layer[69] in a confined pyrrolidinium-based IL.

- **A,H:** Crowded structures formed by cations (A) and anions (H) at high values of κ_{Ion} ($\kappa_{\mathrm{Ion}} > 1.0$) when the net counter charge cannot be provided by a *single* dense layer of counter-ions[18]. We point to the molecular details of this phenomenon for ions with complex molecular geometry like the $[\mathrm{BMIm}]^+$ cations considered in this study: in the molecular simulations a higher counter-charge can be accumulated either by the formation of a distinguishable second sublayer of $[\mathrm{BMIm}]^+$ cations (leading to

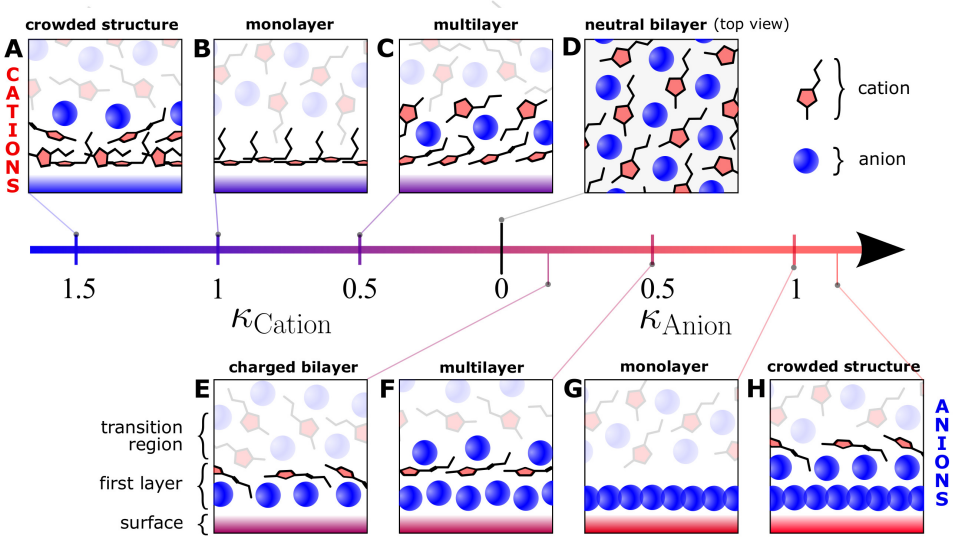


Figure 8: Schematic representation of cations (red pentagons with black tails) and anions (blue balls) packing and orientations close to the electrode (coloured rectangle) at different κ-values. Links between the shown structures and the position on the κ-scale are indicated with lines. In terms of cation and anion packing, the EDL structures formed upon surface charging can be roughly classified as ionic bilayer (D,E), multilayer (C,F), monolayer (B,G) and crowded polyamorphic structures (A,H). The figure is taken from Ref.[9]. Copyright 2014, Elseveir.

crowding) or by reorientation of the $[\mathrm{BMIm}]^+$ rings from parallel to perpendicular orientation relative to the surface (due to electrostriction).

The visualised reorientation of the alkyl chains and the $[\mathrm{BMIm}]^+$ ring (Figure 8B–D) is in agreement with recent spectroscopic studies[43]. Moreover, the molecular representation of the bilayer-to-multilayer restructuring (Figure 3C–F) supports atomic force microscopy (AFM) insights into the EDL structure in ILs[16,35,5,70]. First, using eq. 3 with $\epsilon = 1.6$ we estimate the potential of the monolayer formation of common $[\mathrm{BPyr}]^+$, $[\mathrm{TFSI}]^-$, $[\mathrm{FAP}]^-$ to be -3.1 V, $+2.9$ V and $+2.7$ V, respectively. Consequently, we conclude that the electrode potential range used in experiments[16,35,5,70] corresponds to the κ_{Ion} values between 0 and 0.5–0.7. In both MD simulations and in AFM experiments[16,70], upon $\kappa_{\mathrm{Cation}} \to 0.5$, the number of interfacial layers is increasing (Figure 6); at negative surface charge densities the thickness of the first layer varies due to the reorientation of the cation ring (Figure 8C); and at positive surface charge densities the thickness of the first layer remains constant, although, the alkyl tails may penetrate the anionic sublayer (Figure 8E).

7 Conclusions

The ideas discussed above require further development and an extensive validation by direct experiments and more sophisticated molecular models. However, we note that a number of characteristic structures were observed experimentally at electrochemical conditions

that could serve as some evidences of the monolayer formation. For example, formation of ordered ionic adlayers were observed via scanning tunneling microscopy[71–73]; also, "monolayer" to "bilayer" restructuring induced by confinement was observed by surface forces apparatus[69] and solid-like IL multilayers have been recently reported[74,42].

The results discussed above also prove that the earlier ideas on over-screening to overcrowding[18] and multilayer to monolayer[7,8] structural transitions in ILs obtained by simple models are still generally valid when the molecular structure of IL ions is taken into account.

Overall, we think that molecular simulations can be used for predicting and analysis of ion layering and packing in IL-based interfacial systems. The presented qualitative analysis of interfacial structural transitions in ILs through the κ and λ parameters can be applied for rational design of specific electrochemical and electromechanical applications of ILs[1].

Acknowledgments

We are thankful to Alexei A. Kornyshev and Ruth Lynden-Bell for many fruitful discussions about theoretical and modelling aspects of ILs at CIs. We are grateful for computer time at the EPSRC funded ARCHIE-WeSt High Performance Computer (www.archie-west.ac.uk, EPSRC grant no. EP/K000586/1) and for support from the von Neumann-Institut für Computing, FZ Jülich (Project ID ESMI11).

References

1. Maxim V. Fedorov and Alexei A. Kornyshev. Ionic liquids at electrified interfaces. *Chem. Rev.*, 114(5):2978–3036, 2014.
2. Michel Armand, Frank Endres, Douglas R. MacFarlane, Hiroyuki Ohno, and Bruno Scrosati. Ionic-liquid materials for the electrochemical challenges of the future. *Nat. Mater.*, 8(8):621–629, 2009.
3. Douglas R. MacFarlane, Naoki Tachikawa, Maria Forsyth, Jennifer M. Pringle, Patrick C. Howlett, Gloria D. Elliott, James H. Davis, Masayoshi Watanabe, Patrice Simon, and C. Austen Angell. Energy applications of ionic liquids. *Energy Environ. Sci.*, 7:232–250, 2014.
4. Philippe Hapiot and Corinne Lagrost. Electrochemical reactivity in room-temperature ionic liquids. *Chem. Rev.*, 108:2238–2264, 2008.
5. James Sweeney, Florian Hausen, Robert Hayes, Grant B. Webber, Frank Endres, Mark W. Rutland, Roland Bennewitz, and Rob Atkin. Control of nanoscale friction on gold in an ionic liquid by a potential-dependent ionic lubricant layer. *Phys. Rev. Lett.*, 109(15):155502–5, October 2012.
6. O. Werzer, E. D. Cranston, G. G. Warr, R. Atkin, and M. W. Rutland. Ionic liquid nanotribology: mica-silica interactions in ethylammonium nitrate. *Phys. Chem. Chem. Phys.*, 14(15):5147–5152, 2012.
7. K. Kirchner, T. Kirchner, V. Ivaništšev, and M.V. Fedorov. Electrical double layer in ionic liquids: Structural transitions from multilayer to monolayer structure at the interface. *Electrochim. Acta*, 110:762–771, 2013.

8. Vladislav Ivaništšev and Maxim V. Fedorov. Interfaces between charged surfaces and ionic liquids: Insights from molecular simulations. *Electrochem. Soc. Interface*, 23(1):65–69, 2014.
9. V. Ivaništšev, S. O'Connor, and M. V. Fedorov. Poly(a)morphic portrait of the electrical double layer in ionic liquids. *Electrochem. Commun.*, 48:61–64, November 2014.
10. V. Ivaništšev, K. Kirchner, T. Kirchner, and M. V. Fedorov. Restructuring of the electrical double layer in ionic liquids upon charging. *J. Phys.: Condens. Matter*, 27(10):102101, March 2015.
11. T. Welton. Room-temperature ionic liquids. solvents for synthesis and catalysis. *Chem. Rev.*, 99(8):2071–2083, 1999.
12. J. P. Hallett and T. Welton. Room-temperature ionic liquids: Solvents for synthesis and catalysis. 2. *Chem. Rev.*, 111(5):3508–3576, May 2011.
13. Yaodong Liu, Yi Zhang, Guozhong Wu, and Jun Hu. Coexistence of liquid and solid phases of bmim-PF6 ionic liquid on mica surfaces at room temperature. *J. Am. Chem. Soc.*, 128(23):7456–7457, June 2006.
14. Hua Zhou, Michael Rouha, Guang Feng, Sang Soo Lee, Hugh Docherty, Paul Fenter, Peter T. Cummings, Pasquale F. Fulvio, Sheng Dai, John McDonough, Volker Presser, and Yury Gogotsi. Nanoscale perturbations of room temperature ionic liquid structure at charged and uncharged interfaces. *ACS Nano*, 6(11):9818–9827, October 2012.
15. Florian Buchner, Katrin Forster-Tonigold, Benedikt Uhl, Dorothea Alwast, Nadja Wagner, Hanieh Farkhondeh, Axel Groß, and R. Jürgen Behm. Toward the microscopic identification of anions and cations at the ionic liquid|ag(111) interface: A combined experimental and theoretical investigation. *ACS Nano*, 7(9):7773–7784, September 2013.
16. Hua Li, Ross J. Wood, Frank Endres, and Rob Atkin. Influence of alkyl chain length and anion species on ionic liquid structure at the graphite interface as a function of applied potential. *J. Phys.: Condens. Matter*, 26(28):284115, July 2014.
17. Timo Carstens, René Gustus, Oliver Höfft, Natalia Borisenko, Frank Endres, Hua Li, Ross J. Wood, Alister J. Page, and Rob Atkin. Combined STM, AFM, and DFT study of the highly ordered pyrolytic graphite/1-octyl-3-methyl-imidazolium bis(trifluoromethylsulfonyl)imide interface. *J. Phys. Chem. C*, 118(20):10833–10843, May 2014.
18. M. Z. Bazant, B. D. Storey, and A. A. Kornyshev. Double layer in ionic liquids: Overscreening versus crowding. *Phys. Rev. Lett.*, 106(4):046102–4, 2011.
19. Sami Tazi, Mathieu Salanne, Christian Simon, Pierre Turq, Michael Pounds, and Paul A. Madden. Potential-induced ordering transition of the adsorbed layer at the ionic Liquid/Electrified metal interface. *J. Phys. Chem. B*, 114(25):8453–8459, 2010.
20. Eunsu Paek, Alexander J. Pak, and Gyeong S. Hwang. A computational study of the interfacial structure and capacitance of graphene in [BMIM][PF6] ionic liquid. *J. Electrochem. Soc.*, 160(1):A1–A10, 2013.
21. S. A. Kislenko, R. H. Amirov, and I. S. Samoylov. Molecular dynamics simulation of the electrical double layer in ionic liquids. *J. Phys.: Conf. Ser.*, 418(1):012021–8, 2013.
22. Céline Merlet, Benjamin Rotenberg, Paul A. Madden, and Mathieu Salanne. Computer simulations of ionic liquids at electrochemical interfaces. *Phys. Chem. Chem.*

Phys., 15(38):15781–15792, September 2013.
23. Céline Merlet, David T. Limmer, Mathieu Salanne, René van Roij, Paul A. Madden, David Chandler, and Benjamin Rotenberg. The electric double layer has a life of its own. *J. Phys. Chem. C*, May 2014.
24. Frank Endres, O. Höfft, N. Borisenko, L. H. Gasparotto, A. Prowald, R. Al-Salman, T. Carstens, R. Atkin, A. Bund, and S. Z. El Abedin. Do solvation layers of ionic liquids influence electrochemical reactions? *Phys. Chem. Chem. Phys.*, 12(8):1724–1732, 2010.
25. R. M. Lynden-Bell, A. I. Frolov, and M. V. Fedorov. Electrode screening by ionic liquids. *Phys. Chem. Chem. Phys.*, 14(8):2693–2701, February 2012.
26. V. Ivaništšev, M. V. Fedorov, and R. M. Lynden-Bell. Screening of Ion–Graphene electrode interactions by ionic liquids: The effects of liquid structure. *J. Phys. Chem. C*, 118(11):5841–5847, 2014.
27. Viktoria Nikitina, Sergey A. Kislenko, Renat Ravilevich Nazmutdinov, Michael D. Bronshtein, and Galina Alexandrovna Tsirlina. A Ferrocene/Ferrocenium redox couple at au(111)/Ionic liquid and au(111)/Acetonitrile interfaces: a molecular level view at the elementary act. *J. Phys. Chem. C*, 118(12):6151–6164, January 2014.
28. Svyatoslav Kondrat, Peng Wu, Rui Qiao, and Alexei A. Kornyshev. Accelerating charging dynamics in subnanometre pores. *Nat. Mater.*, 13(4):387–393, April 2014.
29. Hua Li, Mark William Rutland, and Rob Atkin. Ionic liquid lubrication: Influence of ion structure, surface potential and sliding velocity. *Phys. Chem. Chem. Phys.*, 15:14616–14623, 2013.
30. Alexei A. Kornyshev and Rui Qiao. Three-dimensional double layers. *J. Phys. Chem. C*, June 2014.
31. Maxim V. Fedorov and Alexei A. Kornyshev. Ionic liquid near a charged wall: Structure and capacitance of electrical double layer. *J. Phys. Chem. B*, 112(38):11868–11872, 2008.
32. Maxim V. Fedorov and Alexei A. Kornyshev. Towards understanding the structure and capacitance of electrical double layer in ionic liquids. *Electrochim. Acta*, 53(23):6835–6740, 2008.
33. Marcel Drüschler, Natalia Borisenko, Jens Wallauer, Christian Winter, Benedikt Huber, Frank Endres, and Bernhard Roling. New insights into the interface between a single-crystalline metal electrode and an extremely pure ionic liquid: slow interfacial processes and the influence of temperature on interfacial dynamics. *Phys. Chem. Chem. Phys.*, 14(15):5090–5099, 2012.
34. J. Vatamanu, O. Borodin, and G. D. Smith. Molecular insights into the potential and temperature dependences of the differential capacitance of a room-temperature ionic liquid at graphite electrodes. *J. Am. Chem. Soc.*, 132(42):14825–14833, June 2010.
35. Hua Li, Frank Endres, and Rob Atkin. Effect of alkyl chain length and anion species on the interfacial nanostructure of ionic liquids at the au(111)–ionic liquid interface as a function of potential. *Phys. Chem. Chem. Phys.*, 15(35):14624–14633, August 2013.
36. R. Hayes, G. G. Warr, and R. Atkin. At the interface: solvation and designing ionic liquids. *Phys. Chem. Chem. Phys.*, 12(8):1709–1723, 2010.
37. Susan Perkin. Ionic liquids in confined geometries. *Phys. Chem. Chem. Phys.*, 14(15):5052–5062, 2012.

38. Rob Atkin, Natalia Borisenko, Marcel Drüschler, Frank Endres, Robert Hayes, Benedikt Huber, and Bernhard Roling. Structure and dynamics of the interfacial layer between ionic liquids and electrode materials. *J. Mol. Liq.*, 192:44–54, 2014.
39. Markus Mezger, Heiko Schröder, Harald Reichert, Sebastian Schramm, John S. Okasinski, Sebastian Schöder, Veijo Honkimäki, Moshe Deutsch, Benjamin M. Ocko, John Ralston, Michael Rohwerder, Martin Stratmann, and Helmut Dosch. Molecular layering of fluorinated ionic liquids at a charged sapphire (0001) surface. *Science*, 322(5900):424–428, 2008.
40. Yasuyuki Yokota, Tomohiro Harada, and Ken-ichi Fukui. Direct observation of layered structures at ionic liquid/solid interfaces by using frequency-modulation atomic force microscopy. *Chem. Commun.*, 46(45):8627–8629, November 2010.
41. Frank Endres, Natalia Borisenko, Sherif Zein El Abedin, Robert Hayes, and Rob Atkin. The interface ionic liquid(s)/electrode(s): In situ STM and AFM measurements. *Faraday Discuss.*, 154:221–233, 2012.
42. M. Rosa Castillo, José M. Fraile, and José A. Mayoral. Structure and dynamics of 1-butyl-3-methylimidazolium hexafluorophosphate phases on silica and laponite clay: From liquid to solid behavior. *Langmuir*, 28(31):11364–11375, August 2012.
43. Steven Baldelli. Interfacial structure of room-temperature ionic liquids at the Solid–Liquid interface as probed by sum frequency generation spectroscopy. *J. Phys. Chem. Lett.*, 4(2):244–252, 2013.
44. W. Schmickler and E. Santos. *Interfacial Electrochemistry*. Springer, New York, 2 edition, 2010.
45. V. Ivaništšev, K. Kirchner, T. Kirchner, and M.V. Fedorov. Reorganization of the electrical double layer in ionic liquids upon charging. *Phys. Rev. Lett.*, Submitted, 2013.
46. N. Georgi, A. A. Kornyshev, and M. V. Fedorov. The anatomy of the double layer and capacitance in ionic liquids with anisotropic ions: Electrostriction vs. lattice saturation. *J. Electroanal. Chem.*, 649(1-2):261–267, 2010.
47. Stanisław Lamperski, Marcin Waśko, and Douglas Henderson. Solidification of the charged hard-sphere fluid. *Mol. Simulat.*, 39(10):837–841, 2013.
48. Yansen Lauw, Michael D. Horne, Theo Rodopoulos, Vera Lockett, Bulent Akgun, William A. Hamilton, and Andrew R. J. Nelson. Structure of [c4mpyr][NTf2] room-temperature ionic liquid at charged gold interfaces. *Langmuir*, 28(19):7374–7381, 2012.
49. A. A. Kornyshev. Double-layer in ionic liquids: Paradigm change? *J. Phys. Chem. B*, 111(20):5545–5557, 2007.
50. Y. Lauw, M. D. Horne, T. Rodopoulos, and F. A. M. Leermakers. Room-temperature ionic liquids: Excluded volume and ion polarizability effects in the electrical double-layer structure and capacitance. *Phys. Rev. Lett.*, 103(11):117801–4, 2009.
51. Y. Lauw, M. D. Horne, T. Rodopoulos, A. Nelson, and F. A. M. Leermakers. Electrical double-layer capacitance in room temperature ionic liquids: Ion-size and specific adsorption effects. *J. Phys. Chem. B*, 114(34):11149–11154, 2010.
52. Guang Feng, Jingsong Huang, Bobby G. Sumpter, Vincent Meunier, and Rui Qiao. A "counter-charge layer in generalized solvents" framework for electrical double layers in neat and hybrid ionic liquid electrolytes. *Phys. Chem. Chem. Phys.*, 13:14723–14734, 2011.

53. Jianzhong Wu, Tao Jiang, De-en Jiang, Zhehui Jin, and Douglas Henderson. A classical density functional theory for interfacial layering of ionic liquids. *Soft Mat.*, 7(23):11222–11231, 2011.
54. Vincent Démery, David S. Dean, Thomas C. Hammant, Ronald R. Horgan, and Rudolf Podgornik. The one-dimensional coulomb lattice fluid capacitor. *J. Chem. Phys.*, 137(6):064901–16, 2012.
55. Zongzhi Hu, Jenel Vatamanu, Oleg Borodin, and Dmitri Bedrov. A molecular dynamics simulation study of the electric double layer and capacitance of [BMIM][PF6] and [BMIM][BF4] room temperature ionic liquids near charged surfaces. *Phys. Chem. Chem. Phys.*, 15:14234–14247, 2013.
56. George H. Lane. Electrochemical reduction mechanisms and stabilities of some cation types used in ionic liquids and other organic salts. *Electrochim. Acta*, 83:513–528, November 2012.
57. Stephen Fletcher, Victoria Jane Black, Iain Kirkpatrick, and Thomas Stephen Varley. Quantum design of ionic liquids for extreme chemical inertness and a new theory of the glass transition. *J. Solid State Electrochem.*, 17(2):327–337, February 2013.
58. Anabela J. L. Costa, Mário R. C. Soromenho, Karina Shimizu, Isabel M. Marrucho, José M. S. S. Esperança, José N. Canongia Lopes, and Luís Paulo N. Rebelo. Density, thermal expansion and viscosity of cholinium-derived ionic liquids. *ChemPhysChem*, 13(7):1902–1909, 2012.
59. Monika Součková, Jaroslav Klomfar, and Jaroslav Pátek. Surface tension and 0.1 MPa densities of imidazolium-, pyridinium-, pyrrolidinium-, and piperidinium-based tris(pentafluoroethyl)trifluorophosphate ionic liquids. *Fluid Phase Equilibria*, 333:38–46, November 2012.
60. David R. Lide, editor. *CRC Handbook of Chemistry and Physics, 90th Edition*. CRC Press, 90 edition, 2009.
61. B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, 4(3):435–447, 2008.
62. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. A smooth particle mesh ewald method. *J. Chem. Phys.*, 103(19):8577–8593, 1995.
63. In-Chul Yeh and Max L. Berkowitz. Ewald summation for systems with slab geometry. *J. Chem. Phys.*, 111(7):3155–3162, 1999.
64. G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity rescaling. *J. Chem. Phys.*, 126(1):014101–7, 2007.
65. Somisetti V. Sambasivarao and Orlando Acevedo. Development of OPLS-AA force field parameters for 68 unique ionic liquids. *J. Chem. Theory Comput.*, 5(4):1038–1050, April 2009.
66. L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez. PACKMOL: a package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.*, 30(13):2157–2164, 2009.
67. Kathleen Kirchner. *A simulation study of simple ionic liquids near charged walls: The melting of the electric double layer and structural transitions at the interface*. Doctoral thesis, University of Strathclyde, Glasgow, March 2013.
68. Christian Schröder. Collective translational motions and cage relaxations in molecular ionic liquids. 135(2):024502–11, July 2011.

69. Alexander M. Smith, Kevin R. J. Lovelock, Nitya Nand Gosvami, Peter Licence, Andrew Dolan, Tom Welton, and Susan Perkin. Monolayer to bilayer structural transition in confined pyrrolidinium-based ionic liquids. *J. Phys. Chem. Lett.*, 4:378–382, 2013.
70. R. Hayes, N. Borisenko, M. K. Tam, P. C. Howlett, F. Endres, and R. Atkin. Double layer structure of ionic liquids at the au(111) electrode interface: An atomic force microscopy investigation. *J. Phys. Chem. C*, 115(14):6855–6863, 2011.
71. Yuzhuan Su, Jiawei Yan, Miangang Li, Meng Zhang, and Bingwei Mao. Electric double layer of au(100)/Imidazolium-Based ionic liquids interface: Effect of cation size. *J. Phys. Chem. C*, 117(1):205–212, January 2013.
72. Y. Z. Su, Y. C. Fu, J. W. Yan, Z. B. Chen, and B. W. Mao. Double layer of au(100)/Ionic liquid interface and its stability in imidazolium-based ionic liquids. *Angew. Chem. Int. Edit.*, 48(28):5148–5151, 2009.
73. Ge-Bo Pan and Werner Freyland. 2D phase transition of PF6 adlayers at the electrified ionic liquid/Au(111) interface. *Chem. Phys. Lett.*, 427(1-3):96–100, 2006.
74. S. Bovio, A. Podesta, C. Lenardi, and P. Milani. Evidence of extended solidlike layering in [bmim][NTf2] ionic liquid thin films at room-temperature. *J. Phys. Chem. B*, 113(19):6600–6603, May 2009.

Biomolecular Solvation

Matthias Heyden

Max-Planck-Institute für Kohlenforschung, Theoretical Chemistry
Kaiser-Wilhelm-Platz 1, 45470 Mülheim a.d. Ruhr, Germany

E-mail: heyden@kofo.mpg.de

1 Introduction

In this lecture, we discuss applications of molecular dynamics simulations to study biomolecular solvation, i.e. the interactions between biomolecular solutes and their solvent environment. The consequences of these interactions can be studied directly by analyzing the contrast of solvent properties in the bulk and at the biomolecular solute interface. In aqueous solution, the term biological water[1] has been coined for water in the environment of a solvated biomolecule with properties distinguishable from the bulk. More generally, we may refer also to the hydration or solvation shell of a biomolecular solute. The potential relevance of biological water for biomolecular function has been discussed intensely[2,3] in the past decade. Water can be directly involved in the chemistry of an enzymatically catalyzed reaction[4]. Further, protein flexibility is strongly dependent on the presence of hydrating water[5]. Also the protein dynamical transition, the appearance of anharmonic motion in the protein with increasing temperature, was shown to be strongly coupled to the onset of dynamics in the surrounding hydration water[6]. Further, proteins were found to interact strongly with the intermolecular vibrations of the water hydrogen bond network surrounding them[7,8].

In dry or non-aqueous solvation states, proteins are typically kinetically trapped in their respective conformation[9]. In cases where biological function involves any internal motion of the protein, its activity will be significantly affected. The formation of a single layer of hydrating water molecules has been shown to be sufficient to restore dynamical fluctuations of a protein as well as activities[10].

For the study of biological water, simulations offer the advantage of spatial resolution, i.e. the ability to specifically characterize the properties of water molecules in the vicinity of a solvated biomolecule, a specific binding site or catalytic center. In contrast, many experimental techniques have to extract hydration water properties indirectly by comparing a solution to the bulk solvent, e.g. dielectric and far-infrared/terahertz spectroscopy[11,7]. Alternatively, hydrated protein powders are studied in order to study hydration water properties[12]. On the other hand, some experiments offer a certain degree of spatial resolution, such as ultrafast fluorescence[13,14]. Here, dipolar relaxation processes are quantified in the local environment of a fluorescent probe which can be bound covalently to specific residues on the protein surface. However, a correct assessment of the influence of the required probe molecule is not always straight forward. Here, detailed simulations do not only provide crucial information on the properties of the hydration environment, but are also able to play a major role in the interpretation of experimental observations.

In the following section, we will discuss the relevant timescales for biomolecular solvation effects and in particular hydration water. Then, we will describe the microscopic, dynamical processes in water that can be directly analyzed from simulations. Using a simple protein as an example, we will study the sensitivity of hydration water properties to the presence of a biomolecular solute. We will then introduce a methodology to resolve dynamical properties in the three dimensional environment of biomolecular solutes and analyze specifically correlated protein–water vibrational dynamics. Finally, we will describe some useful mathematical and computational concepts, for extensive calculation of auto and cross correlation functions.

Several important topics will not be discussed explicitly, which is why we will briefly mention them here. The field of biomolecular solvation does not just describe the interactions of biomolecular solute molecules with the main solvent, i.e. water. In biologically relevant environments and experimental setups, interactions with co–solutes, co–solvents and ions have a severe impact, for example on protein stabilities and biological activities[15,16].

Furthermore, the native environment of biomolecules is rarely a dilute solution. Typical biomolecular environments, such as the cytoplasm in the interior of the cell, are highly concentrated with water contents below 70%[17]. Under such crowded conditions, the interactions between biomolecules, proteins, nucleic acids, metabolites, play a crucial role and effects the stability of compactly folded proteins, their dynamics and aggregation propensity. Detailed insights into the interactions of biomolecular with their aqueous solvent and their dynamical coupling, can therefore only be regarded as the first step towards understanding complex biomolecules in their even more complex environments.

2 Timescales in Biomolecular Solvation

To the advantage of simulation studies, the timescales required to study the properties of solvents are significantly smaller than for complex biomolecular solutes. For example, water molecules in a room temperature molecular dynamics simulation can be considered as being in equilibrium on timescales of 10–100 ps. On the contrary, a simulated biomolecule is typically captured exclusively in a pseudo–equilibrium state, e.g. its fast fluctuations within one or a few accessible conformations are sampled effectively. Transitions between such conformations occur on microsecond or even millisecond timescales. Therefore, state–of–the–art atomistic simulations, with very few notable exceptions[18], are capable of sampling a few individual conformational transitions at best, but don't even remotely approach equilibrium statistics. Consequently, the properties of water in a biomolecular hydration shell extracted from a simulation are only representative for the pseudo–equilibrium state obtained for the respective solute. The latter is particularly relevant for comparisons to equilibrium properties obtained in experiments.

The concepts described in the following do not only apply to aqueous solvents. Water is the dominant solvent for nucleic acids and globular proteins, whose native fold is even stabilized by water, i.e. the hydrophobicity of buried residues and favorable solvation of exposed sites. Integral membrane proteins experience an entirely different solvating environment. They are primarily embedded in a two-dimensional liquid bilayer of lipids, whose composition is highly variable. Interactions of membrane proteins with water play only a secondary role. However, the interactions of membrane proteins with their lipid environ-

ment will influence the properties of the protein and the lipid solvation layer. Lipid bilayer dynamics occur on significantly slower timescales on the order of tens of nanoseconds. Mixed bilayers formed by more than one lipid constituent exhibit micro– and meso–scopic mixing and demixing processes even on microsecond timescales[19]. Consequently, statistical sampling even of solvent properties is particularly challenging for such systems. Here, we will limit ourselves to a discussion of aqueous solvents for simplicity.

3 Water Dynamics

3.1 Time domain

Dynamical relaxation processes in water are dominated by rearrangements of its tetrahedral hydrogen bond network, which is responsible for a large number of unusual liquid properties[20]. Examples are the wide temperature range for a stable liquid state, the high boiling temperature for such a light molecule and the density maximum at 4°C. Rotational or translational motions of a water molecule require one or more hydrogen bonds formed with neighboring molecules to break. Consequently, this is the rate limiting step, for example for molecular diffusion. In molecular dynamics simulations it is straightforward to characterize water dynamics via the process of breaking and forming hydrogen bonds. For this purpose, a criterion for the definition of an intact hydrogen bond needs to be constructed[21,22]. Such a criterion can be based on the binding energy or a simple geometric definition. The choice of the criterion will affect the quantitative values obtained for the number of hydrogen bonds per water molecule in the bulk liquid and their average lifetime, however qualitative trends, e.g. with temperature or for differences between bulk and hydration shell water, are typically robust. Geometric criteria are commonly used due to their simplicity and independence of the employed energy model, for example the force field. These are usually based on the separation distance between donor (D) and acceptor (A) atoms and the angle formed by the D-H and A-H vectors. Popular choices are D-A separation distances $<$3.5 Å and an angle larger than $>$150°. This criterion, expressed as an operator $\hat{h}$, defines a given hydrogen bond between two molecules as either intact (1) or broken (0)[21,22]. The dynamics of hydrogen bond rearrangements are then conveniently expressed in terms of a time correlation function:

$$C_{\mathrm{HB}}(\tau) = \frac{\langle \hat{h}(t)\hat{h}(t+\tau)\rangle_t}{\langle \hat{h}(t)\hat{h}(t)\rangle_t} \tag{1}$$

Here, the angular brackets $\langle\rangle_t$ denote an ensemble average over all hydrogen bonds and reference times t. The correlation function $C_{\mathrm{HB}}(\tau)$ simply describes the decreasing probability of a hydrogen bond that was found intact at any given time t to be also intact at later times $t+\tau$. This definition also accounts for hydrogen bonds, which were broken intermittently and reform at a later time. A final decay of $C_{\mathrm{HB}}(\tau)$ therefore involves not only breaking of a hydrogen bond, but also a diffusional process which separates the two previously bonded molecules from each other. A typical result for bulk water at room temperature is displayed in Fig. 1A, which shows that the residual probability of an average hydrogen bond to be intact after 10 ps is less than 10%.

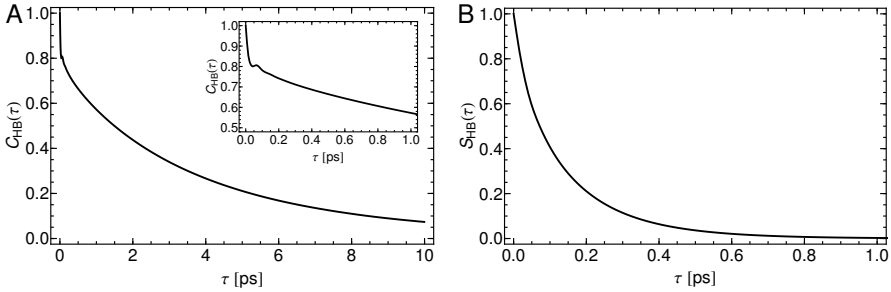


Figure 1: Hydrogen bond correlation functions for bulk water at room temperature. $C_{\mathrm{HB}}(\tau)$ (A) allows for intermittent hydrogen bond breaking events, while $S_{\mathrm{HB}}(\tau)$ does not count hydrogen bonds which experienced any breaking event. The inset in panel A shows the fast initial fluctuations of $C_{\mathrm{HB}}(\tau)$ due to fast hydrogen bond breaking and recovery events on the 100 fs timescale.

A more strict definition of the correlation function utilizes the operator $\hat{H}(t, t+\tau)$, which tests if the given hydrogen bond was broken at any time in the time interval from t to $t+\tau$.

$$S_{\mathrm{HB}}(\tau) = \frac{\langle \hat{h}(t)\hat{H}(t,t+\tau)\rangle_t}{\langle \hat{h}(t)\hat{h}(t)\rangle_t} \tag{2}$$

Intermittent hydrogen bond breaking and immediate recovery events are tightly coupled to molecular vibrations, in particular librational motions (see below), which take place on time scales $<$100 fs. In $C_{\mathrm{HB}}(\tau)$ fluctuations due to these fast events can be observed on this timescale as shown in the inset of Fig. 1A. In practice, $S_{\mathrm{HB}}(\tau)$ therefore depends strongly on the time resolution of the analyzed trajectory and the ability to resolve such transient hydrogen bond breaking events. For converged results, time resolutions on the order of 10 fs are typically required. In Fig; 1B a typical bulk water result is shown (obtained with a time resolution of 8 fs) in which an almost complete decay is observed after 500 fs. While $C_{\mathrm{HB}}(\tau)$ characterizes the more complex dynamics of hydrogen bond network rearrangements and thus follows a multi-exponential decay, $S_{\mathrm{HB}}(\tau)$ provides a direct measure of the hydrogen bonding strength and can be described by a single exponential. Average lifetimes associated with both relaxation processes are either obtained by fitting, integration of the correlation function, e.g. $\tau_{\mathrm{HB}} = \int_0^\infty C_{\mathrm{HB}}(\tau)d\tau$, or simply via $C_{\mathrm{HB}}(\tau_{\mathrm{HB}}) = 1/\mathrm{e}$, in analogy to the time constant of a single exponential decay.

While the correlation functions $C_{\mathrm{HB}}(\tau)$ and $S_{\mathrm{HB}}(\tau)$ are smooth functions describing the average breaking of a hydrogen bond, the actual process of breaking individual hydrogen bonds is discontinuous. Non-transient rupture of a hydrogen bond is typically associated with fast (fs) large jumps of the hydrogen bond angle of up to 60°[23]. These jumps are the result of the fast formation of a new hydrogen bond with an available hydrogen bonding partner in the environment.

As indicated, the decay of $C_{\mathrm{HB}}(\tau)$ is directly related to water molecule rotations and translational diffusion, for which hydrogen bond rupture is a rate–limiting step. To obtain complementary information both processes are often also studied individually. To study the rotational motion of water molecules we define a correlation function based on a vec-

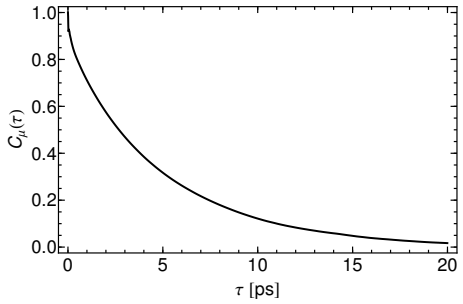


Figure 2: Single molecule dipole correlation function $C_\mu(\tau)$ reporting on rotational relaxations in bulk water at room temperature.

tor defined in the molecule centered coordinate system, for example the molecular dipole moment μ or one of the OH-bond vectors.

For the molecular dipole μ we obtain:

$$C_\mu(\tau) = \frac{\langle \mu(t)\mu(t+\tau)\rangle_t}{\langle \mu(t)\mu(t)\rangle_t} \tag{3}$$

For rigid molecules the expression in Eq. 3 is equivalent to a correlation function of a unit vector along the dipole moment μ, however, for flexible molecules the correlation function also contains fluctuations due to the intramolecular vibrations of the molecule. Essentially the correlation function describes the average cosine of the angle θ between the vector at the reference time t and time $t+\tau$, $C_\mu(\tau) = \cos[\theta(\tau)]$. It decays to 0, when all angles between 0 and 180° are equally probable (averaging to 90°) and all information about the initial orientation of the water molecule is lost as shown in Fig. 2.

Orientational correlations functions are particularly interesting as they are accessible experimentally. For example, femtosecond pump-probe experiments are able to probe the rotation of vibrationally excited O-H bonds[24,25]. Nuclear magnetic relaxation experiments are sensitive to a closely related correlation function, which corresponds to the second order Legendre polynomial of the cosine of θ (while $\cos[\theta(\tau)]$ is equivalent to the first order Legendre polynomial)[26].

Translational diffusion can be easily quantified via the mean squared displacement of the molecular center of mass or atom coordinates:

$$\mathrm{MSD}(\tau) = \langle |\mathbf{r}(t+\tau) - \mathbf{r}(t)|^2 \rangle_t \tag{4}$$

In the limit of long observation times, the slope of the $\mathrm{MSD}(\tau)$ is directly proportional to the diffusion coefficient via the Einstein–Smoluchowski relation:

$$D = \frac{1}{2N_{\mathrm{DIM}}} \lim_{\tau\to\infty} \frac{d\mathrm{MSD}(\tau)}{d\tau} \ , \tag{5}$$

where N_{DIM} corresponds to the number of dimensions in which the motion takes place. A typical mean squared displacement curve is shown for bulk water molecules in Fig. 3.

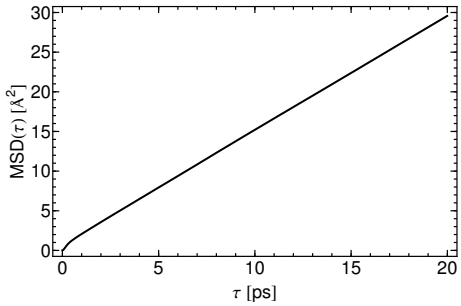


Figure 3: Water molecule mean squared displacements computed from molecular centers of mass reporting on the self–diffusion in bulk water at room temperature.

After a short initial period of ballistic flight (MSD increases quadratically with time, before the tagged water molecule interacts with its neighbors) a transition to a linear evolution is observed after less than 1 ps.

Translational motion of water molecules, i.e. MSD's, also offer a direct, qualitative comparison to experimental observables, in particular from quasi-elastic neutron scattering (QENS)[6]. For more quantitative comparisons the incoherent scattering function can be computed[27].

Diffusion coefficients are equally accessible from velocity time correlation functions.

$$C_{\mathbf{v}}(\tau) = \langle \mathbf{v}(t)\mathbf{v}(t+\tau)\rangle_t \tag{6}$$

This correlation function often decays non–monotonously due to low frequency vibrations in the system as shown for velocity correlation functions of oxygen and hydrogen atoms in Fig. 4. Its decay to zero can be used as a proper equilibration time, i.e. after reinitialization of atomic velocities in a molecular dynamics simulation. The integral over the time correlated motions provide an alternative measure of the self–diffusion coefficient, a Green–Kubo relation.

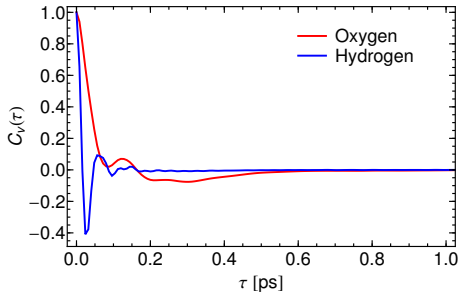


Figure 4: Normalized atomic velocity correlation functions $C_{\mathbf{v}}(\tau)$ for oxygen and hydrogen atoms in bulk water at room temperature.

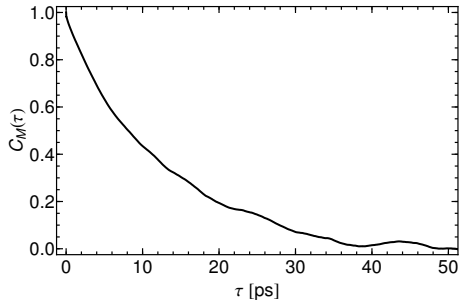


Figure 5: Normalized total dipole correlation function $C_{\mathbf{M}}(\tau)$ describing dielectric relaxation in bulk water at room temperature.

$$D = \frac{1}{N_{\mathrm{DIM}}} \int_{\tau=0}^{\infty} C_{\mathbf{v}}(\tau) d\tau \tag{7}$$

The dynamic properties discussed above analyze the average dynamics of single molecules or hydrogen bonds, and are thus generally referred to as single–particle dynamics. They do not directly contain information on collective dynamics, i.e. the correlated dynamics of interacting molecules.

As an example, the relaxation of the total dipole moment $\mathbf{M}$ of a system does formally contain such information. It is analyzed via a time auto correlation function in analogy to the single molecule dipole moment in Eq. 3 and is displayed for bulk water in Fig. 5.

$$C_{\mathbf{M}}(\tau) = \langle \mathbf{M}(t)\mathbf{M}(t+\tau) \rangle_t \tag{8}$$

The total dipole moment $\mathbf{M}(t)$ can be expressed as the sum of individual molecular dipole moments $\mu(t)$

$$\mathbf{M}(t) = \sum_i \mu_i(t) \tag{9}$$

Substitution of this expression into Eq. 8 leads to a mixed expression of molecule dipole auto correlations and cross correlations between distinct molecules:

$$\langle \mathbf{M}(t)\mathbf{M}(t+\tau) \rangle_t = \langle \sum_i \mu_i(t) \sum_j \mu_j(t+\tau) \rangle_t \tag{10}$$

$$= \sum_i \langle \mu_i(t)\mu_i(t+\tau) \rangle_t + \sum_i \sum_{j \neq i} \langle \mu_i(t)\mu_j(t+\tau) \rangle_t \tag{11}$$

Consequently, the total dipole relaxation formally reports on collective behavior, unless the cross correlation terms are zero (or negligibly small). The relaxation of $C_{\mathbf{M}}(\tau)$ corresponds to the experimental observable of dielectric spectroscopy. Other types of collective dynamics, e.g. the correlated motion of atoms and molecules can be observed in coherent

scattering techniques, which obtain the dynamic structure factor and will be touched upon briefly later[28].

3.2 Frequency domain

Some of the previously discussed correlation functions reveal additional useful information in the frequency domain after Fourier transformation. The total dipole autocorrelation function discussed above is directly related to the infrared spectrum via the fluctuation–dissipation theorem[29]. It relates the equilibrium fluctuations of a given property to the linear response of a system to an external perturbation. In this case the perturbation is the interaction of electromagnetic radiation with the dipole moment of the system, whose equilibrium fluctuations are described in its time correlation function in Eq. 8[30].

$$\alpha(\omega) = F(\omega) \int \exp(\mathrm{i}\omega\tau) \langle \mathbf{M}(t)\mathbf{M}(t+\tau)\rangle_t d\tau \tag{12}$$

The frequency dependent prefactor $F(\omega)$ can be chosen according to various quantum correction schemes, which relate the correlation function observed in classical simulations to the quantum mechanical observable. The so–called harmonic approximation pre–factor $F(\omega) = \frac{1}{4\pi\epsilon_0} \frac{2\pi\omega^2}{3k_\mathrm{B}TVcn(\omega)}$ is recommended in Ref. 32, which can be assumed as proportional to ω^2 if the frequency dependence of the refractive index $n(\omega)$ is neglected. In analogy to Eq. 11 the absorption spectrum can be separated into contributions from single molecule auto correlations, which are equivalent to a Fourier transform of Eq. 3, and cross correlations of distinct molecules.

$$\alpha(\omega) = \quad F(\omega) \int \exp(\mathrm{i}\omega\tau) \sum_i \langle \mu_i(t)\mu_i(t+\tau)\rangle_t d\tau \tag{13}$$

$$+F(\omega) \int \exp(\mathrm{i}\omega\tau) \sum_i \sum_{j\neq i} \langle \mu_i(t)\mu_j(t+\tau)\rangle_t d\tau \tag{14}$$

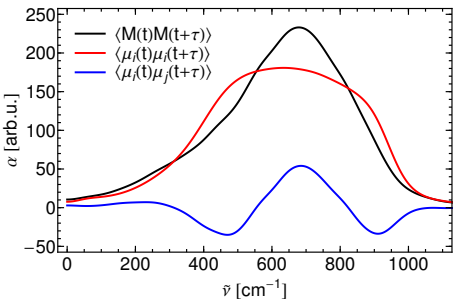


Figure 6: Exemplary absorption spectrum for bulk water obtained from a simulation using a rigid non–polarizable force field model (SPC/E). Note that the experimentally known absorption band at 200 cm^{-1} due to hydrogen bond stretch vibrations[31] is inactive due to the absence of polarization effects. Shown are also spectral contributions from single molecule dipole fluctuations and cross correlations between molecular dipole fluctuations of distinct molecules.

Results obtained from fluctuations of the total dipole moment, as well as contributions to the absorption spectrum from fluctuations of individual molecular dipoles and the cross correlations between them, are shown in Fig. 6 for a simulation of bulk water at room temperature and experimental density using the non–polarizable SPC/E force field model. Water molecules in this model are described as rigid and intramolecular vibrations are not existing. Therefore we restrict the analysis to frequencies below $1000\,\mathrm{cm}^{-1}$ where intermolecular vibrations are observed. It should be noted that the absorption spectrum obtained from simulations of non–polarizable water models describe the experimentally observed spectrum rather poorly in the presented formalism[33–35]. Intermolecular vibrations lead to mutual dynamic polarizations of interacting water molecules, which affect dipole fluctuations and thus the IR activity of the respective vibrational modes. This is particularly notable for the experimentally determined absorption band of hydrogen bond stretch vibrations, which take place at roughly $200\,\mathrm{cm}^{-1}$[36]. In the non–polarizable water model this vibration is not IR–active and therefore this band is absent in the spectrum in Fig. 6. The main feature which can be observed in the total IR spectrum in Fig. 6 (black) is the so–called libration band, which is very broad and ranges from $300\,\mathrm{cm}^{-1}$ to $1000\,\mathrm{cm}^{-1}$. The underlying vibrations are fast rotational motions of water molecules which are hindered by the hydrogen bonds formed with neighboring molecules. These motions are mainly responsible for the transient hydrogen bond breaking events, which lead to the decay of $S_{\mathrm{HB}}(\tau)$ introduced earlier. Water molecule rotations involve a substantial change of the dipole moment, which results in the IR activity of the vibrations. The broad range of frequencies is due to the three different moments of inertia for the independent rotational axes of a water molecule, as well as the heterogeneities of the individual hydrogen bonding environments of each molecule.

This description of the librational absorption band already indicates, that single molecule dipole fluctuations dominate the total IR signal. This is demonstrated in Fig. 6 (red). Only the shape of the total absorption band is modified by the correlated molecular dipole fluctuations (blue). The latter describe the response of the environment to the librational motion of an individual water molecule. Negative intensities indicate that the total dipole moment change due to librations of a single molecule are partially compensated by a correlated response in the environment. Again, we note that the non–polarizable model used here only provides a poor representation of correlated dipole moments in water due to the missing dynamic polarization effects. See Ref. 35 for a detailed analysis of dipolar correlations and a comparison between non–polarizable force field and *ab initio* simulations that capture dynamic polarization effects.

The velocity auto correlation function $C_{\mathbf{v}}(\tau)$ contains direct information on the fluctuations of the atomic motion in the system, i.e. its vibrations. Its Fourier transform reports on the vibrational motion in the system, irrespective of the cross section of these vibrations with spectroscopic measurements, such as infrared activity. For non–polarizable simulation models, it thus provides a significantly more realistic description of the vibrational spectrum than the infrared absorption computed directly from total dipole fluctuations[35]. Using proper mass weighting, e.g. by scaling of atomic velocities with the square root of the atomic mass, $\tilde{\mathbf{v}}_i = \sqrt{m_i}\mathbf{v}_i$ the vibrational density of states $I(\omega)$ is obtained, which describes the distribution of the kinetic energy in the system over vibrational frequencies.

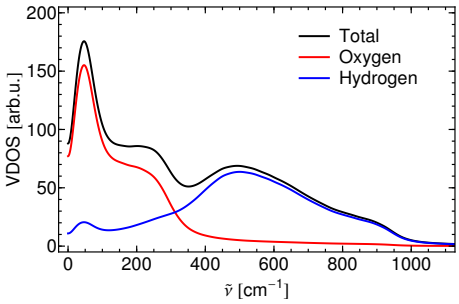


Figure 7: Vibrational density of states of water at room temperature with separate contributions from oxygen and hydrogen atoms. While frequencies below 200 cm^{-1} are dominated by oxygens, higher frequency vibrations are attributed primarily to vibrational motion of the lighter hydrogen atoms.

$$I(\omega) = \frac{1}{3k_{\mathrm{B}}T} \int \exp(\mathrm{i}\omega\tau) \langle \tilde{\mathbf{v}}(t) \tilde{\mathbf{v}}(t+\tau) \rangle_t d\tau \tag{15}$$

Figure 7 displays the vibrational density of states (VDOS) computed separately from oxygen and hydrogen velocities, as well as the sum of both. Comparison to the IR spectrum in Fig. 6 shows, that the IR active libration band is dominated by hydrogen motion, which matches the expectation for water molecule rotations around the molecule center of mass. On the other hand vibrations at frequencies below 300 cm^{-1} are dominated by oxygen motion, e.g. vibrations of the molecular center of mass. Clearly visible is now the vibrational band due to hydrogen bond stretch motions as a pronounced shoulder at 200 cm^{-1}, which has been absent in the absorption spectrum in Fig. 6. Below 100 cm^{-1}, an additional vibration band is observed, which is commonly assigned to bending vibrations involving the angle of two hydrogen bonds with a common center[36]. Recently, it has been shown that the underlying motion is likely to be more complex involving at least as many as eight water molecules forming a local hydrogen bonded cluster of molecules in the liquid[37].

4 Hydration Water of Biomolecules

4.1 Structure and density fluctuations

The hydration water of biomolecules exhibits distinct properties from bulk water due to its interactions with atoms on the protein surface[1]. The tetrahedral structure of the water hydrogen bond network is responsible for a low density of the liquid. Neutron and x-ray scattering experiments indicate that defects in this tetrahedral structure at biomolecule–water interfaces can lead to a partial collapse of the hydrogen bond network structure, resulting in an increased water density[38]. Simulations show, that a substantial amount of this increase (two thirds) can be attributed to a geometrical effect of defining any random surface in the water hydrogen bond network, which is not related to actual structural changes in the hydration shell[39]. However, an approximately 5% increase in the hydration water

density (here for the lysozyme protein) can be attributed to actual changes in the hydrogen bond network structure, i.e. a reduced oxygen-oxygen distance and an increased average coordination number of each water molecule.

Attractive interactions with charged groups on protein surfaces are reported to have a specific effect on the increased water density in its direct vicinity[39]. However, topographical and geometrical properties of the protein surface are often found to be important factors for structural changes in hydration water, which do not necessarily correlate with chemical properties of the protein–water interface.

Apart from this, hydrophobic surface patches of biomolecules do not offer hydrogen bond donor or acceptor sites to hydrating water molecules. Consequently, hydrating water molecules try to optimize hydrogen bonding with other water molecules in order to avoid loose dangling OH bonds. This restricts the number of hydrogen bonding configurations for hydrating water molecules, resulting in an entropic cost. This is the basis for the popular 'iceberg' model[40], i.e. clathrate–like structures in the hydration water of hydrophobic solutes. For extended hydrophobic surfaces, this enthalpic cost can be too great, resulting in the loss of hydrogen bonds. This can be detected spectroscopically due to the distinct O-H vibration frequency of dangling O-H bonds[41], which point to the hydrophobic surface. In this case, enthalpic costs for hydrophobic hydration become significant[42].

A specific property of water hydrating hydrophobic solutes are significantly increased density fluctuations[43]. The absence of strong solute–solvent interactions facilitates cavity formations, which lead to the observed fluctuations. Density fluctuations are tightly coupled to the compressibility, which can be several times higher at an extended hydrophobic surface than in the bulk. Single hydrophilic groups within a hydrophobic surface patch can already be sufficient to bind hydrating water molecules, effectively suppressing the fluctuations[44]. Consequently, the magnitude of density fluctuations in the vicinity of a hydrophobic surface are context dependent, e.g. depend on the properties of nearby surfaces. The characteristic density fluctuations of water hydrating hydrophobic surfaces bears potential application for a microscopic hydrophobicity scale[45]. One of the key steps for protein aggregation or ligand binding is the dewetting of the binding interface. Hence, enhanced density fluctuations in the vicinity of hydrophobic residues are crucial for the kinetics of binding events[46].

Structural changes in biomolecular hydration water, e.g. increased densities, are restricted to one layer of water molecules, which are in direct contact to functional groups on the protein surface[39]. Apart from the density, the orientation of water molecules can be perturbed on significantly larger length scales. In the first few hydration layers water molecules preferentially orient themselves according to the electric field generated by surface groups of the protein. At larger distances to the protein, however, interactions with the very large dipole moment of many proteins (often 100 Debye and more) can become dominant, resulting in a preferred average orientation of water molecules in the hydration shell[47]. Such effects have been proposed to be very long-ranged (nanometer length scales), but may be reduced substantially by the presence of ions, which affect the dielectric properties of the solution.

4.2 Hydration water dynamics

Apart from the previously discussed density fluctuations, also the dynamical processes of the water hydrogen bond network react to the presence of a biomolecular solute sur-

face. Solute–water hydrogen bonds can be stronger than water–water hydrogen bonds, in particular for charged carboxyl or amino groups. In addition, the kinetics of hydrogen bond rearrangements are significantly perturbed by the solute surface. This can be accurately captured by taking into account the reduced volume accessible to competing water molecules to attack a solute–water hydrogen bond, which is necessary to form the transition state of hydrogen bond rearrangement events. Both factors combined are found to be sufficient to describe increased lifetimes of solute–water hydrogen bonds[48].

The decreased mobility of water molecules directly bonded to polar groups on a biomolecular surface results in a retardation of water–water hydrogen bond rearrangement dynamics in the first hydration shell, and to limited degree also higher hydration shells. This is shown in Fig. 8, where water-water hydrogen bond correlation functions are shown for water molecules selected in successive hydration layers based on the distance to the closest atom of a simulated protein (here the DNA binding domain of the λ-repressor)[49]. Hydrogen bonds within the selected set of water molecules are then followed for 10 ps to obtain a rough spatial resolution of water dynamics in the vicinity of the protein. Water–water hydrogen bond lifetimes defined simply as $C_{\mathrm{HB}}(\tau_{\mathrm{HB}} = 1/\mathrm{e}$ are found to be increased by more than 50% in the direct vicinity of the protein solute. With increasing distance to the protein the lifetimes approach the bulk value, which is recovered at approximately 6 Å distance.

Water-water hydrogen bond dynamics are heterogeneously distributed on the surface of the protein. Due to the absence of protein–water hydrogen bonds for hydrophobic sidechains, one would assume a weaker retardation in their vicinity, as is observed for example in the vicinity of hydrophobic model solutes[50]. However, as discussed for observations of modified protein hydration water densities, geometrical and topological factors contribute as well for complex protein solute surfaces. With hydrophobic residues being in average less exposed to the solvent, e.g. within crevices of the protein surface, geometrical constraints for such water molecules can be increased. As a result, hydrogen bond rearrangement dynamics can be found to be even slower than for solvent exposed hydrophilic sidechains. This is reported for example for the λ-repressor DNA binding domain[49].

The retardation of hydrogen bond dynamics is largely reflected by other dynamical pro-

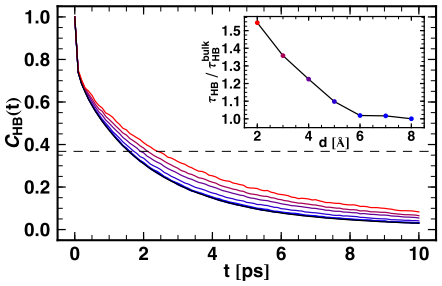


Figure 8: Hydrogen bond correlation functions $C_{\mathrm{HB}}(t)$ for hydrogen bonds between water molecules in the hydration shell of the λ-repressor DNA binding domain[49]. Shown are correlation functions for water molecules selected at the beginning of 10 ps time windows in successive shells of 4 Å thickness. The inset displays the associated relaxation times τ_{HB} normalized by the bulk value as a function of average distance to the closest protein atom. The bulk water correlation function is shown in black.

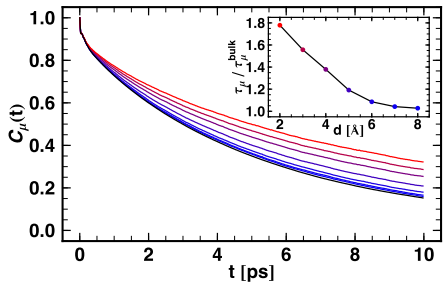


Figure 9: Single molecule dipole correlation functions $C_\mu(t)$ for water molecules in the hydration shell of the λ-repressor DNA binding domain[49]. Shown are correlation functions for water molecules selected at the beginning of 10 ps time windows in successive shells of 4 Å thickness. The inset displays the associated relaxation times τ_μ normalized by the bulk value as a function of average distance to the closest protein atom. The bulk water correlation function is shown in black.

cesses, such as molecular rotation[51,52,49]. In Fig. 9 the single molecule dipole correlation function is shown for the same selections of water molecules used in Fig. 8. The associated lifetimes show that the orientational relaxation is even decreased by up to 80% in the direct vicinity of the protein surface. Bulk behavior is roughly restored at distances of 6–7 Å distance to the protein, comparable to the hydrogen bond rearrangement dynamics.

Similar observations are made for self–diffusion processes[53,49]. The $\mathrm{MSD}(t)$ of the same selections of water molecules are displayed in Fig. 10. The self–diffusion coefficient in the first hydration layer is reduced to just 55% of the bulk value. However, it seems that self–diffusion is not yet fully restored at distances of up to 8 Å from the protein. It needs to be taken into consideration that the translational motion in the vicinity of the protein surface is restricted basically to two dimensions. Neglecting the reduced dimensionality is Eq. 4 thus implicitly reduces the obtained self–diffusion coefficients as a consequence of the geometrical constraints experienced by hydrating water molecules, in addition to the retardation of water dynamics due to interactions with the protein surface.. However, the

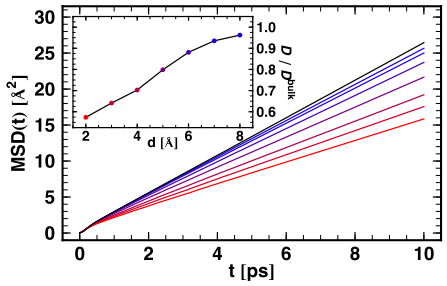


Figure 10: Mean squared displacements for water molecules in the hydration shell of the λ-repressor DNA binding domain[49]. Shown are time evolutions of the $\mathrm{MSD}(t)$ for water molecules selected at the beginning of 10 ps time windows in successive shells of 4 Å thickness. The inset displays the associated diffusion constants normalized by the bulk value as a function of average distance to the closest protein atom. The bulk water $\mathrm{MSD}(t)$ is shown in black.

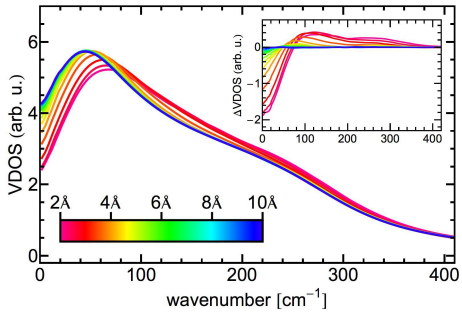


Figure 11: Water oxygen VDOS in the hydration shell of the λ-repressor DNA binding domain. Shown are the Fourier transforms of atomic velocity time auto correlation functions with increasing distance from the closest protein atom. The inset shows the differences to the bulk water oxygen VDOS.

result in Fig. 10 indicates that solute–induced retardation effects on hydration water self–diffusion have a slightly longer range than for hydrogen bond dynamics and rotational motion.

Figure 11 displays the evolution of the water oxygen VDOS in the hydration shell of the protein form 0 to 400 cm^{-1}. The observed bands below 100 cm^{-1} and at roughly 200 cm^{-1} correspond to hydrogen bond bending and stretching vibrations discussed in the context of Fig. 7[54,49]. Figure 11 shows that the retardation of water dynamics also affects the intermolecular vibrations. It is reflected by a blue–shift of vibrational frequencies, which is particularly noticeable for the hydrogen bond bending mode, whose peak shifts from roughly 45 cm^{-1} in the bulk to 65 cm^{-1} directly at the protein surface. To a lesser extent this is also observed for the hydrogen bond stretch shoulder at 200 cm^{-1}. Blue–shifted vibrational frequencies can be interpreted as the result of a larger effective force constant. This finding is consistent with the longer lifetime of water–water hydrogen bonds in the protein hydration shell. While the nature of the water–water interaction potential will not be affected in the protein hydration water (at least in a non–polarizable potential), the long–lived vibrating hydrogen bonds spend more time in the minimum of the highly anharmonic potential due to the retarded rearrangement dynamics. In the vicinity of this minimum, the curvature, i.e. the effective force constant, is larger, leading to blue–shifted vibrational frequencies.

5 Hydration Water Properties and Correlated Dynamics in 3D

In the previous section, the properties of water in the hydration shell of a protein have been analyzed using a simple distance selection criterion at the beginning of a relatively short time window to obtain limited spatial resolution. This time window (e.g. 10 ps) has been only partially sufficient to analyze, for example, the single molecule dipole correlation function, which only decayed to values between 0.15 and 0.35 (Fig. 9). Nevertheless, the $\mathrm{MSD}(\tau = 10\,ps)$ corresponds to values of 16 to 25 Å^2 corresponding to average displacements of 4 to 5 Å, thus limiting substantially the spatial resolution that can be obtained.

For a spatially resolved analysis of dynamical water properties in three dimensions we

thus have to focus on processes which are fast enough for a meaningful local description, i.e. significantly faster than translational diffusion. Further, a distance based selection criterion which follows individual molecules for short time windows is not very practical in this case. For this purpose, we utilize the local density of a given atomic or molecular property X. With the spatial coordinates of particle i at time t given as $\mathbf{r}_i(t)$ it is defined as:

$$\rho_X(\mathbf{r}, t) = \sum_i X_i(t)\delta\left[\mathbf{r}_i(t) - \mathbf{r}\right] \tag{16}$$

This expression of the density takes into account that local time dependent fluctuations of the property X at coordinates $\mathbf{r}$ are not necessarily attributed to a single molecule. However, the infinitely sharp δ function is not useful for practical purposes as the probability to find a particle at any given coordinate $\mathbf{r}$ is zero. Instead, we approximate the density by substituting the δ-functions with a three–dimensional Gaussian kernel to obtain a smooth local density expression[37,35,55,56]:

$$\rho_X(\mathbf{r}, t) \approx \sum_i X_i(t) \frac{1}{(2\pi\sigma^2)^{3/2}} \exp\left[-\frac{|\mathbf{r}_i(t) - \mathbf{r}|^2}{2\sigma^2}\right] \tag{17}$$

The width σ of the kernel determines the spatial resolution in a controlled manner. It also determines the timescale of processes that can be fully observed when a molecule travels through the observation volume. Here, we aim to resolve the vibrational motion of water molecules, i.e. the vibrational density of states. For this purpose, we compute the smooth local density of scaled atomic velocities $\tilde{\mathbf{v}}(t)$. Intermolecular vibrations in water occur on a sub–picosecond timescale and we want to ensure that at any given time only one molecule contributes to the density sampled at $\mathbf{r}$. For this purpose, we typically choose σ between 0.4 and 0.5 Å[37]. The corresponding volume is small enough to ensure that only a single water molecule dominates the velocity density signal at any given time, allowing the description of single particle vibrations. Second, the typical residence time of a molecule is large enough to observe the intermolecular vibrations.

The time dependent smooth local density can be conveniently computed, for example, on a space filling grid allowing to analyze its time dependent fluctuations anywhere in the simulated system[56]. The Fourier transform of the time auto correlation function of the velocity density $\rho_{\tilde{\mathbf{v}}}(\mathbf{r}, t)$ provides a spatially resolved expression for the vibrational density of states.

$$I(\mathbf{r}, \omega) = \frac{1}{3k_B T} \int \exp(\mathrm{i}\omega\tau) \langle \rho_{\tilde{\mathbf{v}}}(\mathbf{r}, t) \rho_{\tilde{\mathbf{v}}}(\mathbf{r}, t+\tau) \rangle_t d\tau \tag{18}$$

The same approach can be employed using the water molecular dipole moment μ, allowing to compute spatially resolved contributions of single molecule dipole fluctuations to the infrared absorption spectrum (within the limitations of the simulation model).

$$\alpha(\mathbf{r}, \omega) = F(\omega) \int \exp(\mathrm{i}\omega\tau) \langle \rho_\mu(\mathbf{r}, t) \rho_\mu(\mathbf{r}, t+\tau) \rangle_t d\tau \tag{19}$$

However, the great advantage of the proper localization of time dependent local fluctuations on a spatial grid is the ability to compute spatially resolved cross correlations between them. For this purpose, correlation functions between the fluctuations sampled at two grid points A and B with coordinates $\mathbf{r}_A$ and $\mathbf{r}_B$ are computed. For the case $A = B$, this corresponds to local self–correlations, i.e. single particle dynamics. In case $A \neq B$, we obtain information on correlated or collective dynamics at both grid points. Due to the fixed coordinates $\mathbf{r}_A$ and $\mathbf{r}_B$, these cross correlations can be analyzed as a function of the separation vector $\Delta\mathbf{r} = \mathbf{r}_B - \mathbf{r}_A$ or the corresponding distance $|\Delta\mathbf{r}|$.

$$C_{\rho,X}(\mathbf{r}_A, \mathbf{r}_B, \tau) = \langle \rho_X(\mathbf{r}_A, t)\rho_X(\mathbf{r}_B, t+\tau)\rangle_t \tag{20}$$

This has been performed for an extensive *ab initio* simulation of bulk water using both, smooth local densities of molecular dipoles μ and scaled atomic velocities $\tilde{\mathbf{v}}$[37,35]. The molecular dipole approach allowed to study the contributions of cross correlated molecular dipoles to the absorption spectrum with spatial resolution, i.e. as a function of separation distance as well as full resolution of cross correlations in the environment of a water molecule. Further, the same approach based on atomic velocities revealed the extent of correlated vibrational motion in the water hydrogen bond network.

The latter will be employed here to analyze correlations between protein vibrations and hydration water vibrations[55,56]. We will focus on the low–frequency vibrations in the water hydrogen bond network, which allows us to neglect contributions from hydrogen atoms. A smooth local velocity density is computed either on a space filling three dimensional grid[56], or on selected points covering the surface of a simulated solvated protein[55] (here the λ-repressor DNA binding domain, which was used as an example also in the previous section). Here, we will employ a three dimensional grid, which is defined in a protein centered coordinate system and follows its translational and rotational motion. The time dependent transformation, which allows the analysis in this coordinate system can be obtained from a minimization of the root mean squared deviations of the protein coordinates. Its rotational part is also applied to the atomic velocities to be able to follow velocity fluctuations along chosen directions in this coordinate system, as discussed below. Thus, we obtain spatial resolution of water oxygen velocity fluctuations, and therefore of the hydrogen bond network vibrations, in the environment of the protein. For a stable fold of the protein, its non–hydrogen atoms are found to oscillate around their average position. For them, no further localization is therefore required.

For every point of the space–filling three dimensional grid, we now select the closest non–hydrogen atom of the protein surface, neglecting small fluctuations of protein atom coordinates, which would render this assignment time dependent. We now compute the time cross correlation function between water oxygen velocity density fluctuations $\rho_{\tilde{\mathbf{v}},O}$ at each individual grid point and the velocities of the closest protein surface atom $\tilde{\mathbf{v}}_P$. The Fourier transform of this cross correlation function now contains information on the correlated vibrational motion of the protein surface and its hydration water hydrogen bond network[56].

$$I(\mathbf{r}, \omega) = \int \exp(\mathrm{i}\omega\tau)\langle \tilde{\mathbf{v}}_P(t)\rho_{\tilde{\mathbf{v}},O}(\mathbf{r}, t+\tau)\rangle_t d\tau \tag{21}$$

For the further analysis, we now select specifically the longitudinal correlated vibra-

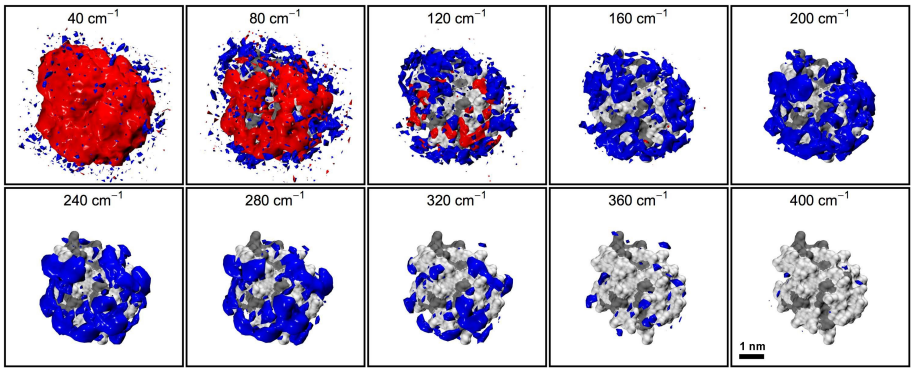


Figure 12: Isosurfaces indicate volume elements with increased negative (blue) and positive (red) intensities of the longitudinal components of Fourier transformed cross correlation functions of protein surface atom velocities and water oxygen velocity densities $I(\mathbf{r},\omega)^{\parallel}$ according to Eq. 22 at various frequencies spanning the spectral range of intermolecular vibrations in the water hydrogen bond network[56]. The protein surface is colored according to the presence of hydrophobic sidechains (dark) or polar sidechains and backbone moieties (light).

tions, i.e. the fluctuations of protein atom and water oxygen velocity densities along the vector connecting the average coordinate of the protein atom and the current grid point in the protein centered coordinate system $\Delta\mathbf{r} = \mathbf{r} - \mathbf{r}_P$. $\Delta\mathbf{r}$ simultaneously corresponds to the normal vector of the respective protein surface, thus an expression equivalent to Eq. 21 provides information on longitudinal vibrations extending from the protein surface into the hydration shell.

$$I(\mathbf{r},\omega)^{\parallel} = \int \exp(\mathrm{i}\omega\tau)\langle \left[\tilde{\mathbf{v}}_P(t)\frac{\Delta\mathbf{r}}{|\Delta\mathbf{r}|}\right]\left[\rho_{\tilde{\mathbf{v}},O}(\mathbf{r},t+\tau)\frac{\Delta\mathbf{r}}{|\Delta\mathbf{r}|}\right]\rangle_t d\tau \tag{22}$$

The result is high dimensional as we obtain resolution in space and frequency. A visualization is provided in Fig. 12. At various frequencies spanning the spectral range of intermolecular vibrations of the water hydrogen bond network, volume elements with either increased negative (blue) and positive (red) cross correlations are represented.

The results shows that correlated vibrational motion of the protein surface and its hydration shell is not homogeneously distributed in the protein environment, but depends on the chemical details of the protein-water interface and the protein vibrational motion in which these surface are involved. Further, correlations tend to be long–ranged for low frequencies below 100 cm^{-1} and restricted to the first hydration layer for higher frequencies. Positive correlations dominate the lowest frequencies. They correspond to the concerted motion of protein and water atoms along the same direction, for example translation of the whole protein and its hydration shell. Higher frequencies are dominated by negative intensities, which correspond to vibrational motion of protein and water atoms in opposite directions, which is a characteristic of stretch vibrations. Specifically notable are these negative intensities at frequencies beyond 200 cm^{-1}, which are observed to be localized at specific sites of the protein–water interface. These correspond to stretch vibrations of protein–water hydrogen bonds, which are found in the vicinity of polar surface groups. High intensities at high frequencies beyond 300 cm^{-1} can be associated to charged surface

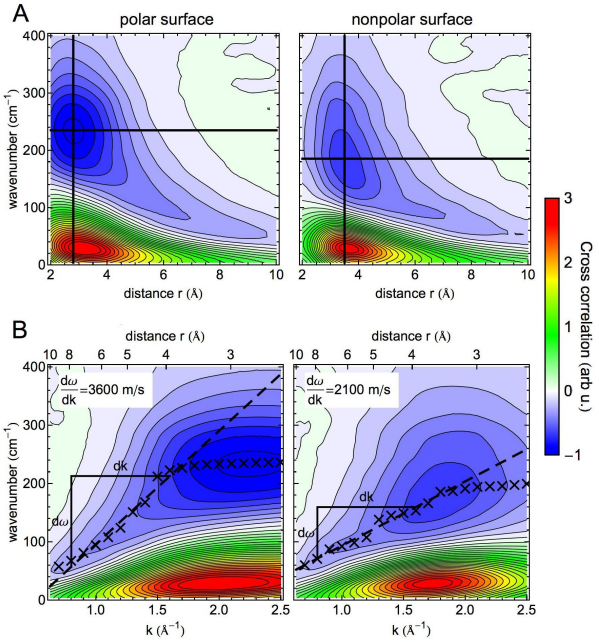


Figure 13: Surface averaged spectra of the longitudinal cross correlation components $I(\mathbf{r},\omega)^{\parallel}$ as a function of the distance r from the protein surface (A)[56]. Results are shown for polar (left panel) and non–polar (right panel) surfaces of the protein. Crossing black lines indicate the maximum negative intensity, which reports on stretch vibrations of protein–water interactions. In panel (B) the same data is shown as a function of the reciprocal distance, $k = 2\pi/r$. The corresponding real–space distances are indicated on the alternative x–axis. The dispersive behavior of the negative intensity peak in the range of intermolecular distances from 0.8 Å^{-1} to 1.6 Å^{-1} is described by a linear fit. Its slope determines the propagation velocity of associated collective motions also indicated in the plots.

groups, which form strong non–covalent bonds with hydrating water molecules.

Using the three–dimensional data presented in Fig. 12, we can average the correlation intensities over polar and non–polar surfaces of the protein (see Ref. 56 for details) as a function of distance from the protein surface[55,56].

The crossed lines in panel A of Fig. 13 indicate the peak position of the negative cross correlations, which stem from the stretch vibrations of protein–water interactions, in particular protein–water hydrogen bonds in the case of the polar protein surface (left panel). Vibrations of water molecules against aliphatic moieties are primarily observed for non–polar protein surfaces (right panel), which is reflected by the decreased average frequency. While these (anti–)correlations in the first hydration shell (within 3 Å of the protein surface) are implicit by the underlying intermolecular stretch vibration, it can be observed that these correlations persist to substantially larger distances up to 10 Å for the protein. The frequency decreases with increasing distance. For the analysis of this distance dependent frequency of the correlation intensity peak, we refer to panel B of Fig. 13, which displays the same data as a function of the reciprocal distance $k = 2\pi/r$ [55]. In this representation, we can observe that the intensity peak of the negative correlations follow a dispersive

behavior, i.e. their frequency changes linearly with k, in the intermolecular range corresponding to distances beyond 4 Å, i.e. k<1.6 $Å^{-1}$. This feature is characteristic for a propagating collective motion, essentially a sound wave. Their propagation velocity is given by the slope of the dispersion curve and is indicated for the polar and non–polar surface hydration water in Fig. 13B[28,55]. Similar collective modes are well–known for bulk water and are termed fast sound for the microscopic, intermolecular length scales observed here. The name is attributed to the significantly lower propagation velocity of macroscopic sound in water, which is observed for much larger distances, i.e. smaller k[57]. The presence of propagating collective motion in the hydration water hydrogen bond network is not surprising by itself. However, it is notable that the properties of modes extending from the protein surface are strongly influenced by the direct protein–water vibrations and therefore chemical details of the protein surface. This is reflected by the considerably lower propagation velocity of the longitudinal collective modes extending from the non–polar protein surface into the hydration water environment.

The spatially resolved smooth local density of water oxygen velocities $\rho_{\tilde{\mathbf{v}}}(\mathbf{r}, t)$ describes in principle a mass current density. The longitudinal spectrum of these mass current density cross correlations $I(k, \omega)^{\parallel}$ displayed in Fig. 13B is therefore equivalent to longitudinal current spectra obtained in coherent scattering experiments. It is directly related to the dynamic structure factor $S(k, t)$ via $I(k, \omega)^{\parallel} = \frac{\omega^2}{k^2} S(k, \omega)$[35]. However, the computational approach described here allows a real space resolution of the underlying correlated vibrations and collective phenomena, which cannot be obtained in experiments. However, the qualitative picture of distinct fast sound propagation velocities in the vicinity of polar and non–polar solute moieties has been observed experimentally for water in concentrated solutions of hydrophobic and hydrophilic peptide solutions[58].

6 Efficient Computation of Correlation Functions

The methods discussed here for the computational analysis of biomolecular hydration water properties rely extensively on calculations of auto and cross correlation functions of time dependent fluctuations. Therefore we use this opportunity to discuss efficient computational strategies for the implementation of such calculations.

The equation for the time auto correlation function of time dependent fluctuation $X(t)$ and $Y(t)$ looks deceptively simple:

$$C(\tau) = \langle X(t) Y(t + \tau) \rangle_t \tag{23}$$

Hence, it seems straight forward to compute the products, for example scalar products if the X and Y represent vectors, for the values of X and Y at equal times for $C(\tau = 0)$, and for subsequent delay times τ to obtain $C(\tau)$. However, in particular if large data sets are to be analyzed or a large number of correlation functions needs to be computed, this simple operation can become computationally quite demanding or even prohibitive. For example, the three–dimensional distribution of frequency dependent cross correlation intensities sampled on a three dimensional grid shown in Fig. 12, requires the calculation of 32768 cross correlation functions, each for data sets containing 12500 time points. The straight forward calculation of the time correlation functions scales approximately with the number of different delay times τ multiplied by the number of available time points

N. Given that the desired number of delay times may be a substantial fraction of available time points N at which the fluctuations of $X(t)$ and $Y(t)$ have been sampled, this results in a scaling of the order of $O(N^2)$.

Regarding the size of the data set in the described example, this would cause a substantial computational effort. It is simple to parallelize, but more elegant solutions to the problem exist. These become even more efficient for larger data sets or correlations in multiple dimensions. The calculation of the Fourier transform of the time correlation function of the fluctuations $X(t)$ and $Y(t)$:

$$C_{X,Y}(\omega) = \int \exp(\mathrm{i}\omega\tau)\langle X(t)Y(t+\tau)\rangle_t d\tau \tag{24}$$

is substantially simpler than calculating the correlation directly. It can be computed as a simple product from the Fourier transformed signals $X(\omega)$ and $Y(\omega)$.

$$C_{X,Y}(\omega) = X(\omega)\overline{Y(\omega)} \ , \tag{25}$$

where overlining indicates the complex conjugate. For the common case of a real–valued time dependent signal $Y(t)$, the complex conjugate of its Fourier transform simply corresponds to a switched sign in the argument of the plane wave expression[56].

$$\overline{Y(\omega)} = \int \exp(-\mathrm{i}\omega t)Y(t)dt \tag{26}$$

If the signals X and Y are identical, $C_{X,X}(\omega)$ corresponds to the Fourier transformed of an auto correlation function. Auto correlations are by definition symmetric, i.e. $C_{X,X}(\tau) = C_{X,X}(-\tau)$. Its Fourier transform therefore contains only real coefficients, i.e. it corresponds to a cosine transformation. Therefore $C_{X,X}(\omega)$ takes on only real values.

If the signals X and Y are not identical, we are computing a cross correlation function. In this case, the Fourier transformed cross correlation function $C_{X,Y}(\omega)$ is complex valued, unless both time dependent signals $X(t)$ and $Y(t)$ are symmetric in time, resulting in real Fourier transforms $X(\omega)$ and $Y(\omega)$. If $X(t)$ and $Y(t)$ are time dependent fluctuations sampled in a molecular dynamics simulation, they are usually not symmetric in time.

Another special case is provided by cross correlation functions computed from fluctuations under equilibrium conditions[56]. In a true equilibrium, the direction of time is arbitrary. In this situation, the cross correlation function of the equilibrium fluctuations $X(t)$ and $Y(t)$ are symmetric in time as well. In numerical simulations, this is not exactly fulfilled due to the finite sampling, although in the limit of good statistics the imaginary part of $C_{X,Y}(\omega)$ will be very small and can be neglected. This neglect of the small imaginary components is equivalent to a symmetrization operation applied on $C_{X,Y}(\tau)$ in the time domain to enforce proper equilibrium behavior.

The straightforward calculation of the Fourier transform of the signals $X(t)$ and $Y(t)$, as well as its inverse to obtain the time correlation function, also scales with $O(N^2)$. Since multiple Fourier transforms are involved, it may seem that the procedure is even more computationally demanding than the straight forward time domain calculation. However, the numerical Fourier transformation involves a large number of identical terms. The fast Fourier transform (FFT) algorithm takes advantage of this, thus reducing the computational

cost to the order of $O(N\log N)$. The resulting speed–up of the calculation is substantial, in particular for large data sets.

The concept is easily extended to correlations in multiple dimensions. For example, in Ref. 37 the smooth local density of water molecular dipoles and atomic velocities is sampled for a bulk water system on a cubic grid, which filled the entire simulation cell. The cross correlations for each pair of grid points were computed and resolved as a function of the displacement vector $\Delta\mathbf{r}$ between two grid points. This corresponds to a correlation function in time and space.

$$C_{\rho\rho}(\tau, \Delta\mathbf{r}) = \langle \rho(t, \mathbf{r})\rho(t + \tau, \mathbf{r} + \Delta\mathbf{r})\rangle_{t,\mathbf{r}} \tag{27}$$

Periodic boundary conditions were applied in the simulation, consequently the sampling grid is intrinsically periodic as well. Therefore the distribution of the time dependent fluctuations of the smooth local density of molecular dipoles μ or scaled atomic velocities $\tilde{\mathbf{v}}$, ρ, can be accurately represented in reciprocal space, i.e. by the spatial Fourier transform.

For time dependent fluctuations sampled in a finite time window, we imply periodicity in time by using the Fourier transform, although this is normally not the case. However, for long term sampling of fluctuations around an equilibrium average, the introduced error can be neglected. For the spatial distribution, which is sampled with much fewer points in each dimension (32 in Ref. 37), actual periodicity is however advisable.

In correspondence to Eq. 25 the space–time Fourier transform

$$C_{\rho\rho}(\omega, \mathbf{k}) = \int\int \exp(\mathrm{i}\omega\tau)\exp(\mathrm{i}\mathbf{k}\Delta\mathbf{r})C_{\rho\rho}(\tau, \Delta\mathbf{r})d\tau\, d^3\Delta\mathbf{r} \tag{28}$$

can be directly computed from the space–time Fourier transform of the smooth local density ρ.

$$C_{\rho\rho}(\omega, \mathbf{k}) = \rho(\omega, \mathbf{k})\overline{\rho(\omega, \mathbf{k})} \tag{29}$$

Inverse Fourier transformation then leads to the real space and time domain expression. Each inverse transformation can be carried out independently of each other to obtain the desired quantity.

$$C_{\rho\rho}(\tau, \Delta\mathbf{r}) = \int\int \exp(-\mathrm{i}\omega\tau)\exp(-\mathrm{i}\mathbf{k}\Delta\mathbf{r})C_{\rho\rho}(\omega, \mathbf{k})d\omega\, d^3\mathbf{k} \tag{30}$$

$$C_{\rho\rho}(\omega, \Delta\mathbf{r}) = \int \exp(-\mathrm{i}\mathbf{k}\Delta\mathbf{r})C_{\rho\rho}(\omega, \mathbf{k})d^3\mathbf{k} \tag{31}$$

$$C_{\rho\rho}(\tau, \mathbf{k}) = \int \exp(-\mathrm{i}\omega\tau)C_{\rho\rho}(\omega, \mathbf{k})d\omega \tag{32}$$

Acknowledgments

The by author is grateful for financial support by the Cluster of Excellence RESOLV (EXC 1069) funded by the Deutsche Forschungsgemeinschaft.

References

1. S.-K. Pal, J. Peon, B. Bagchi, and A.-H. Zewail. Biological water: Femtosecond dynamics of macromolecular hydration. *J. Phys. Chem. B*, 106(48):12376–12395, 2002.
2. M. Chaplin. Do we underestimate the importance of water in cell biology? *Nature Rev. Mol. Cell Biol.*, 7:861–866, 2006.
3. P. Ball. Water as an active constituent in cell biology. *Chem. Rev.*, 108(1):74–108, 2008.
4. M.-W. Makinen, W. Maret, and M.-B. Yim. Neutral metal-bound water is the base catalyst in liver alcohol dehydrogenase. *Proc. Natl. Acad. Sci. U.S.A*, 80(9):2584–2588, 1983.
5. J. Knab, J.-Y. Chen, and A. Markelz. Hydration dependence of conformational dielectric relaxation of lysozyme. *Biophys. J.*, 90:2576–2581, 2006.
6. K. Wood, M. Plazanet, F. Gabel, B. Kessler, D. Oesterhelt, D.-J. Tobias, G. Zaccai, and M. Weik. Coupling of protein and hydration-water dynamics in biological membranes. *Proc. Natl. Acad. Sci. U.S.A*, 104(46):18049–18054, 2007.
7. S. Ebbinghaus, S.-J. Kim, M. Heyden, X. Yu, U. Heugen, M. Gruebele, D.-M. Leitner, and M. Havenith. An extended dynamical hydration shell around proteins. *Proc. Natl. Acad. Sci. U.S.A*, 104(52):20749–20752, 2007.
8. M. Grossman, B. Born, M. Heyden, D. Tworowski, G. B. Fields, I. Sagi, and M. Havenith. Correlated structural kinetics and retarded solvent dynamics at the metalloprotease active site. *Nat. Struct. Mol. Biol.*, 18(10), 2011.
9. A.-M. Klibanov. Improving enzymes by using them in organic solvents. *Nature*, 409:241–246, 2001.
10. A. Zaks and A. M. Klibanov. The effect of water on enzyme action in organic media. *J. Biol. Chem.*, 263(17):8017–21, 1988.
11. A. Oleinikova, P. Sasisanker, and H. Weingärtner. What can really be learned from dielectric spectroscopy of protein solutions? a case study of ribonuclease a. *J. Phys. Chem. B*, 108(24):8467–8474, 2004.
12. J.H. Roh, J.E. Curtis, S. Azzam, V.N. Novikov, I. Peral, Z. Chowdhuri, R.B. Gregory, and A.P. Sokolov. Influence of hydration on the dynamics of lysozyme. *Biophys. J.*, 91(7):2573 – 2588, 2006.
13. S.-K. Pal, J. Peon, and A.-H. Zewail. Ultrafast surface hydration dynamics and expression of protein functionality: α–chymotrypsin. *Proc. Natl. Acad. Sci. U.S.A*, 99(24):15297-15302, 2002.
14. L. Zhang, L. Wang, Y.-T. Kao, W. Qiu, Y. Yang, O. Okobiah, and D. Zhong. Mapping hydration dynamics around a protein surface. *Proc. Natl. Acad. Sci. U.S.A*, 104(47):18461–18466, 2007.
15. S. N. Timasheff. Water as ligand: preferential binding and exclusion of denaturants in protein unfolding. *Biochemistry*, 31(41):9857–9864, 1992.
16. S. N. Timasheff. Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proc. Natl. Acad. Sci. U.S.A*, 99(15):9721–9726, 2002.
17. R. J. Ellis and A. P. Minton. Cell biology: join the crowd. *Nature*, 425(6953):27, 2003.

18. D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, and W. Wriggers. Atomic-level characterization of the structural dynamics of proteins. *Science*, 330(6002):341–346, 2010.
19. G. Khelashvili, A. Grossfield, S. E. Feller, M. C. Pitman, and H. Weinstein. Structural and dynamic effects of cholesterol at preferred sites of interaction with rhodopsin identified from microsecond length molecular dynamics simulations. *Proteins Struct. Funct. Bioinf.*, 76(2):403–417, 2009.
20. P. Ball. *H_2O: A biography of water*. Weidenfeld & Nicolson, Orion House, London, 1999.
21. A. Luzar and D. Chandler. Hydrogen-bond kinetics in liquid water. *Nature*, 379:55–57, 1996.
22. B. Bagchi. Water dynamics in the hydration layer around proteins and micelles. *Chem. Rev. Lett.*, 105(9):3197–3219, 2005.
23. D. Laage and J. T. Hynes. A molecular jump mechanism of water reorientation. *Science*, 311(5762):832–835, 2006.
24. D.-E. Moilanen, E.-E. Fenn, Y.-S. Lin, J.-L. Skinner, B. Bagchi, and M.-D. Fayer. Water inertial reorientation: Hydrogen bond strength and the angular potential. *Proc. Natl. Acad. Sci. U.S.A*, 105(14):5295–5300, 2008.
25. K. J. Tielrooij, N. Garcia-Araez, M. Bonn, and H. J. Bakker. Cooperativity in ion hydration. *Science*, 328(5981):1006–1009, 2010.
26. V.-P. Denisov and B. Halle. Protein hydration dynamics in aqueous solution. *Faraday Discuss.*, 103:227–244, 1996.
27. M. Tarek and D.-J. Tobias. The dynamics of protein hydration water: A quantitative comparison of molecular dynamics simulations and neutron-scattering experiments. *Biophys. J.*, 79(6):3244–3257, 2000.
28. M. Tarek and D.-J. Tobias. Single-particle and collective dynamics of protein hydration water: A molecular dynamics study. *Phys. Rev. Lett.*, 89(27):275501, 2002.
29. M. E. Tuckerman. *Statistical mechanics: Theory and Molecular Simulation*. Oxford University Press Inc., New York, New York, USA, 2010.
30. D.-A. McQuarrie. *Statistical mechanics*. University Science Books, Sausalito, CA 94965, USA, 2000.
31. J.-E. Bertie and Z. Lan. Infrared intensities of liquids: The intensity of the oh stretching band of liquid water revisited, and the best current values of the optical constants of H_2O(l) at 25 °C between 15,000 and 1 cm^{-1}. *Appl. Spec.*, 50(8):1047–1057, 1996.
32. R. Ramírez, T. López-Ciudad, P. Kumar, and D. Marx. Quantum corrections to classical time-correlation functions: Hydrogen bonding and anharmonic floppy modes. *J. Chem. Phys.*, 121(9):3973–3983, 2004.
33. P.-A. Madden and R.-W. Impey. On the infrared and Raman spectra of water in the region 5-250 cm^{-1}. *Chem. Phys. Lett.*, 123(6):502–506, 1986.
34. P.-L. Silvestrelli, M. Bernasconi, and M. Parrinello. Ab initio infrared spectrum of liquid water. *Chem. Phys. Lett.*, 277:478–482, 1997.
35. M. Heyden, J. Sun, H. Forbert, G. Mathias, M. Havenith, and D. Marx. Understanding the origins of dipolar couplings and correlated motion in the vibrational spectrum of water. *J. Phys. Chem. Lett.*, 3(16):2135–2140, 2012.

36. G.-E. Walrafen. Raman spectrum of water: Transverse and longitudinal acoustic modes below $\approx$ 300 cm^{-1} and optic modes above $\approx$ 300 cm^{-1}. *J. Phys. Chem.*, 94(6):2237–2239, 1990.
37. M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith, and D. Marx. Dissecting the THz spectrum of liquid water from first principles *v*ia correlations in time and space. *Proc. Natl. Acad. Sci. U.S.A.*, 107(27):12068–12073, 2010.
38. D.-I. Svergun, S. Richard, M.-H.-J. Koch, Z. Sayers, S. Kuprin, and G. Zaccai. Protein hydration in solution: Experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci. U.S.A*, 95(5):2267–2272, 1998.
39. F. Merzel and J.-C. Smith. Is the first hydration shell of lysozyme of higher density than bulk water? *Proc. Natl. Acad. Sci. U.S.A*, 99(8):5378–5383, 2002.
40. H. S. Frank and M. W. Evans. Free volume and entropy in condensed systems iii. entropy in binary liquid mixtures; partial molal entropy in dilute solutions; structure and thermodynamics in aqueous electrolytes. *J. Chem. Phys.*, 13:507–532, 1945.
41. P.-N. Perera, K.-R. Fega, C. Lawrence, E.-J. Sundstrom, J. Tomlinson-Phillips, and Dor Ben-Amotz. Observation of water dangling OH bonds around dissolved nonpolar groups. *Proc. Natl. Acad. Sci. U.S.A*, 106:12230–12234, 2009.
42. Y.-K. Cheng and P. J. Rossky. Surface topography dependence of biomolecular hydrophobic hydration. *Nature*, 392(6677):696–699, 1998.
43. A. J. Patel, P. Varilly, and D. Chandler. Fluctuations of water near extended hydrophobic and hydrophilic surfaces. *J. Phys. Chem. B*, 114(4):1632–1637, 2010.
44. A. P. Willard and D. Chandler. Coarse-grained modeling of the interface between water and heterogeneous surfaces. *Faraday Discuss.*, 141:209–20, 2008.
45. S. N. Jamadagni, R. Godawat, and S. Garde. Hydrophobicity of proteins and interfaces: insights from density fluctuations. *Ann. Rev. Chem. Biomol. Eng.*, 2:147, 2011.
46. D. Chandler. Interfaces and the driving force of hydrophobic assembly. *Nature*, 437:640–647, 2005.
47. D. R. Martin and D. Matyushov. Hydration shells of proteins probed by depolarized light scattering and dielectric spectroscopy: Orientational structure is significant, positional structure is not. *J. Chem. Phys.*, 141(22):22D501, 2014.
48. F. Sterpone, G. Stirnemann, J.-T. Hynes, and D. Laage. Water hydrogen-bond dynamics around amino acids: The key role of hydrophilic hydrogen-bond acceptor groups. *J. Phys. Chem. B*, 114(5):2083–2089, 2010.
49. M. Heyden and M. Havenith. Combining THz spectroscopy and md simulations to study protein-hydration coupling. *Methods*, 52:74–83, 2010.
50. G. Niehues, M. Heyden, D. A. Schmidt, and M. Havenith. Exploring hydrophobicity by thz absorption spectroscopy of solvated amino acids. *Faraday Discuss.*, 150:193, 2011.
51. R. Abseher, H. Schreiber, and O. Steinhauser. The influence of a protein on water dynamics in its vicinity investigated by molecular dynamics simulation. *Proteins Struct. Funct. Genet.*, 25:366–378, 1996.
52. A.-R. Bizzarri and S. Cannistraro. Molecular dynamics of water at the protein-solvent interface. *J. Phys. Chem. B*, 106:6617–6633, 2002.
53. V.-A. Makarov, M. Feig, B.-K. Andrews, and B.-M. Pettitt. Diffusion of solvent around biomolecular solutes: A molecular dynamics simulation study. *Biophys. J.*,

75:150–158, 1998.
54. S. Chakraborty, S.-K. Sinha, and S. Bandyopadhyay. Low-frequency vibrational spectrum of water in the hydration layer of a protein: A molecular dynamics simulation study. *J. Phys. Chem. B*, 111(48):13626–13631, 2007.
55. M. Heyden and D. J. Tobias. Spatial dependence of protein-water collective hydrogen-bond dynamics. *Phys. Rev. Lett.*, 111(21):218101, 2013.
56. M. Heyden. Resolving anisotropic distributions of correlated vibrational motion in protein hydration water. *J. Chem. Phys.*, 141(22):22D509, 2014.
57. F. Sciortino and S. Sastry. Sound propagation in liquid water: The puzzle continues. *J. Chem. Phys.*, 100:3881–3893, 1994.
58. D. Russo, A. Orecchini, A. De Francesco, F. Formisano, A. Laloni, C. Petrillo, and F. Sacchetti. Brillouin neutron spectroscopy as a probe to investigate collective density fluctuations in biomolecules hydration water. *Spectrosc. Int. J.*, 27(5-6):293–305, 2012.

Simulation Techniques for Solvation-Induced Surface-Interactions at Prescribed Water Chemical Potential

Alexander Schlaich[1], Bartosz Kowalik[1], Matej Kanduč[1], Emanuel Schneck[2], and Roland R. Netz[1]

[1] Fachbereich Physik, Freie Universität Berlin, Germany

E-mail: rnetz@physik.fu-berlin.de

[2] Biomaterials Department, Max-Planck Institute of Colloids and Interfaces, Potsdam, Germany

In computer simulations, accounting for the solvent's chemical potential is of great importance when the studied system is in equilibrium with an external solvent reservoir. We first review several techniques that enable determining the chemical potential of small molecules in computer simulations. Their advantages and disadvantages are discussed via numerous numerical analyses with a particular emphasis on water as the universal biological solvent. We then review in detail the recently introduced thermodynamic extrapolation technique, which allows to precisely quantify the interaction between surfaces at prescribed solvent chemical potential. The technique is illustrated for various model surfaces interacting across water. The interaction is found to exhibit qualitatively different behavior between soft and stiff surfaces due to layering effects. We furthermore discuss the importance of surface compressibility, van der Waals forces and the cavitation transition. Finally, we demonstrate how atomistic computer simulations enable reproducing the experimentally observed hydration repulsion between phospholipid membranes, which is crucial for a deeper understanding of the underlying repulsion mechanisms on an atomic scale.

1 Introduction

The importance of understanding interactions between molecules that build biological systems as well as complexes that appear in the synthesis, design, and manipulation of materials at the nanometer scale was recognized from the very beginning of nanoscience. It has been realized that the behavior of all these systems is governed by fundamental molecular interactions such as electrostatic, van der Waals, and hydration interactions[1]. The understanding of interactions between solvated surfaces is of outstanding importance for industrial applications, e.g. for lubrication, or in the analysis of many chemical reactions. Moreover, the sub-micrometer structural organization of all living organisms is to a great extent determined by the physical interactions between the surfaces of the organism's functional units.

In biological systems, water plays the role of the universal solvent and has hence been at the focus of many theoretical and computational studies. As any other liquid, it spatially redistributes in thermodynamic equilibrium, such that its chemical potential is uniform. The various complications when enforcing constant chemical potentials in computer simulations are the main topic of this review. In the next section, we present different techniques to determine the chemical potential of small molecules. We then focus on chemical potentials in one- and two-phase systems (Sec. 3). In Section 4, we describe a method to perform simulations at prescribed solvent chemical potential between two interacting

surfaces at varying separations. In the last section, we employ this method to calculate the hydration pressure between different surfaces.

2 Measuring Chemical Potentials in Computer Simulations

Over the last decades, the research field dealing with molecular interactions witnessed vast intensity of research and progress. Due to the enormous complexity with many molecular details relevant at nanometer scales in soft-matter and biological systems, the research nowadays increasingly relies on computer simulations. One way of computer implementations are Molecular Dynamics (MD) simulations, where all interacting particles (atoms, molecules) are treated classically. In these simulations, the dynamics of the system and the temporal evolution of the quantities are obtained by a numerical integration of Newton's equations of motion[2,3]. Most of what we discuss in the following also applies to coarse-grained methods, where groups of atoms and molecules are lumped into particles with appropriate interactions. Simulations can be performed in several different ensembles, the most commonly used being the NVT and the NpT ensembles in which the particle numbers, the temperature, and either the volume or the pressure are held constant. In both ensembles, the temperature is held constant by a thermostat, which scales the velocity of the particles such that the correct Maxwell-Boltzmann distribution is maintained. In order to maintain the prescribed pressure in the NpT ensemble, the simulation box size is continuously adjusted. For both the thermostat and the pressure coupling there are many different implementations, which we will not discuss here. We refer the interested reader to the literature[4–6].

As hydration is essential for all biological systems, modeling of water is an important task in simulations. Many water models have been developed, among those the Simple Point Charge (SPC) model and its modification, the Simple Point Charge/Extended (SPC/E) model[7]. In these models, each water molecule is represented by three interaction sites: The positive electric charges are fixed to the two hydrogen sites, whereas a negative charge that exactly balances those resides on the oxygen. These partial charges give rise to the water dipole moment and mimic the formation of hydrogen bonds. Additionally, the oxygen atoms in water molecules interact between each other via the Lennard-Jones (LJ) interaction, accounting for both dispersion attraction and Pauli repulsion. Formally, the interaction between two water molecules a and b then is defined as

$$W_{ab} = \sum_{i}^{\text{on } a} \sum_{j}^{\text{on } b} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} + \frac{C_{12}}{r_{\text{OO}}^{12}} - \frac{C_6}{r_{\text{OO}}^{6}}, \tag{1}$$

where the two sums run over all three sites in a molecule. Here, q_i and q_j are partial charges, r_{ij} the corresponding distance between the interacting sites, and C_{12} as well as C_6 the LJ coefficients that act between two oxygen atoms r_{OO} apart. Also other, more complex, water models have been developed, like TIP4P and TIP5P, which use four and five interaction sites, respectively. A comparison of various properties of these models can be found in Ref.[8]. Computationally more demanding water models account also for the polarizability, which in principle can yield a more accurate description of water properties. Due to its simplicity, robustness, and wide-spread use, we will focus on the SPC/E model in the following, but invite the interested reader to read more about water modeling in Ref.[9].

As mentioned above, in computer simulations the number of particles is usually kept fixed. However, for inhomogeneous multi-component systems partitioning effects become important, which requires accounting for the chemical potential of the individual species. In fact, chemical potentials are important for many aspects of physical chemistry, as particles move from regions of higher chemical potential to the region of lower chemical potential upon equilibration. For simulations of systems that are in equilibrium with a bulk solvent, it therefore needs to be ensured that the solvent chemical potential in the simulated section matches that of the bulk. In the following we will review some basic methodology to determine chemical potentials of small molecules in computer simulations.

2.1 Test particle insertion

Test particle insertion (TPI)[10] is an efficient method to calculate the chemical potential of solute or solvent molecules in computer simulations. The main idea is to compute the chemical potential by measuring the average interaction energy of a particle which is randomly inserted into the system. That way, an existing trajectory of a simulation can be used and only the excess energy needs to be evaluated. The major advantages are the simplicity and robustness, as well as low computational costs. However, at high particle densities, problems arise due to the overlap of molecules[11,12], which we will discuss for the case of water simulations in Section 3.3. Here, we derive mathematical expressions for TPI in the NVT and NpT ensembles[2].

In a system of several different particle species $\alpha, \beta, \gamma, \ldots$ the chemical potential μ_α of a particular species α is defined as the derivative of the free energy change with respect to the respective particle number N_α,

$$\begin{aligned}\mu_\alpha &= \left(\frac{\partial G}{\partial N_\alpha}\right)_{p,T,N_{\beta\neq\alpha}} \\ &= \left(\frac{\partial F}{\partial N_\alpha}\right)_{V,T,N_{\beta\neq\alpha}} .\end{aligned} \tag{2}$$

Here, we introduced the Gibbs free energy G and the Helmholtz free energy F, which are the relevant thermodynamic potentials in NpT and NVT ensembles, respectively. In the following, we will focus on systems consisting of only one species of particles and skip the index α, however, the calculation is analogous for many-component systems. We will also assume our system to be enclosed in a rectangular simulation box and we normalize all coordinates r_i^j by the box vector $\mathbf{L}$ with $V = L^x L^y L^z$ being the volume of our system, i.e. we obtain the normalized coordinates of the particles $s_i^j = r_i^j / L^j$, where the upper index refers to the spatial direction. The partition function in the NVT ensemble then reads as

$$Q(N,V,T) = \frac{V^N}{\Lambda^{3N} N!} \int_0^1 \mathrm{d}^3\mathbf{s}_1 \ldots \int_0^1 \mathrm{d}^3\mathbf{s}_N \mathrm{e}^{-\beta U_N(\mathbf{s}_1,\ldots,\mathbf{s}_N;\mathbf{L})}, \tag{3}$$

and in the NpT ensemble as

$$Q(N,p,T) = \int \mathrm{d}V \frac{V^N}{\Lambda^{3N} N!} \int_0^1 \mathrm{d}^3\mathbf{s}_1 \ldots \int_0^1 \mathrm{d}^3\mathbf{s}_N \mathrm{e}^{-\beta[U_N(\mathbf{s}_1,\ldots,\mathbf{s}_N;\mathbf{L})+pV]}, \tag{4}$$

where U_N represents the system Hamiltonian of N particles. We now make use of the fact that the free energies are given by the logarithm of the partition functions,

$$\mu = \frac{\partial G}{\partial N} \approx -k_\mathrm{B}T\left(G(N+1) - G(N)\right) = -k_\mathrm{B}T\,\ln\left(\frac{Q_{N+1}}{Q_N}\right), \tag{5}$$

where we approximated the derivative by a difference, which is valid for systems with a sufficiently large particle number N. In the NVT ensemble, this yields

$$\begin{aligned}
\mu_\mathrm{NVT} &= -k_\mathrm{B}T\,\ln\left(\frac{\frac{V^{N+1}}{\Lambda^{3(N+1)}(N+1)!}\prod_{k=1}^{N+1}\int_0^1 \mathrm{d}^3\mathbf{s}_k\,\mathrm{e}^{-\beta U_{N+1}}}{\frac{V^N}{\Lambda^{3N}N!}\prod_{k=1}^{N}\int_0^1 \mathrm{d}^3\mathbf{s}_k\,\mathrm{e}^{-\beta U_N}}\right)\\
&= -k_\mathrm{B}T\,\ln\left(\frac{V}{\Lambda^3(N+1)}\int_0^1 \mathrm{d}^3\mathbf{s}_{N+1}\frac{\prod_{k=1}^{N}\int_0^1 \mathrm{d}^3\mathbf{s}_k\,\mathrm{e}^{-\beta U_N}\mathrm{e}^{-\beta\Delta U}}{\prod_{k=1}^{N}\int_0^1 \mathrm{d}^3\mathbf{s}_k\,\mathrm{e}^{-\beta U_N}}\right)\\
&= k_\mathrm{B}T\,\ln\left(\frac{(N+1)\Lambda^3}{V}\right) - k_\mathrm{B}T\,\ln\left(\int_0^1 \mathrm{d}^3\mathbf{s}_{N+1}\left\langle \mathrm{e}^{-\beta\Delta U}\right\rangle_N\right).
\end{aligned} \tag{6}$$

In this calculation, we split the total Hamiltonian U_{N+1} into a part U_N containing the interactions between the first N particles and a part $\Delta U = U_{N+1} - U_N$ that accounts for the interactions of the $(N+1)$-th particle with all other particles. The brackets $\langle ... \rangle_N$ denote the canonical ensemble average. Similarly, the chemical potential in the NpT ensemble is given by

$$\begin{aligned}
\mu_\mathrm{NpT} &= -k_\mathrm{B}T\,\ln\left(\frac{\int \mathrm{d}V\,\frac{V^{N+1}}{\Lambda^{3(N+1)}(N+1)!}\,\mathrm{e}^{-\beta pV}\prod_{k=1}^{N+1}\int_0^1 \mathrm{d}^3\mathbf{s}_k\,\mathrm{e}^{-\beta U_{N+1}}}{\int \mathrm{d}V\,\frac{V^N}{\Lambda^{3N}N!}\,\mathrm{e}^{-\beta pV}\prod_{k=1}^{N}\int_0^1 \mathrm{d}^3\mathbf{s}_k\,\mathrm{e}^{-\beta U_N}}\right)\\
&= -k_\mathrm{B}T\,\ln\left(\frac{\langle V\rangle}{\Lambda^3(N+1)}\int_0^1 \mathrm{d}^3\mathbf{s}_{N+1}\frac{\int \mathrm{d}V\,\mathrm{e}^{-\beta pV}V^N\prod_{k=1}^{N}\int_0^1 \mathrm{d}^3\mathbf{s}_k\,\mathrm{e}^{-\beta U_N}V\,\mathrm{e}^{-\beta\Delta U}}{\langle V\rangle\int \mathrm{d}V\,\mathrm{e}^{-\beta pV}V^N\prod_{k=1}^{N}\int_0^1 \mathrm{d}^3\mathbf{s}_k\,\mathrm{e}^{-\beta U_N}}\right)\\
&= k_\mathrm{B}T\,\ln\left(\frac{(N+1)\Lambda^3}{\langle V\rangle}\right) - k_\mathrm{B}T\,\ln\left(\int_0^1 \mathrm{d}^3\mathbf{s}_{N+1}\frac{\left\langle V\mathrm{e}^{-\beta\Delta U}\right\rangle_{N,V}}{\langle V\rangle}\right).
\end{aligned} \tag{7}$$

In contrast to the NVT ensemble, we are averaging over $V\mathrm{e}^{-\beta\Delta U}$ rather than $\mathrm{e}^{-\beta\Delta U}$, which means that we also take into account the fluctuations of the volume of our system. In Eq. (7) we see that if $V = \text{const.}$, we arrive again at the expression for the NVT ensemble. In the following, we therefore drop the index denoting the ensembles and use Eq. (7) for both ensembles.

In Eqs. (6) and (7), we identify two contributions to the chemical potential: The first term, depending only on the particle number density, is the ideal gas chemical potential,

and the second is usually referred to as the excess potential depending on the specific interactions. Thus we write the total chemical potential as

$$\mu = \mu^{id} + \mu^{ex}. \tag{8}$$

The TPI method can now be used to obtain a numerical value for the excess contribution, which equals the difference in the chemical potential of a non-interacting ideal gas particle and a real particle interacting with the rest of the system. At this point it should be noted that the TPI method can be used in both Monte Carlo as well as in Molecular Dynamics simulations, since only a trajectory containing N particles is needed that sufficiently samples the phase space spanned by all coordinates (and the volume for the NpT ensemble). For the application in MD simulations it is crucial to realize that the ensemble average involving N particle positions and the integration over the particle position s_{N+1} are interchangeable in their order, as they are independent of each other. Having simulated an N-particle system, we extract a large number of uncorrelated configurations from it, choose a random position at which we insert an additional randomly oriented test particle and measure its interaction energy ΔU with the entire system. The insertions are only virtual, meaning only an energy measurement is undertaken without a permanent addition of the particle to the trajectory. From each individual energy difference, the Boltzmann factor $\mathrm{e}^{-\beta\Delta U}$ is computed and averaged over all insertion positions of the test particle and all configurations.

Some additional comments have to be made on the averaging: In configurations extracted from Monte Carlo and MD simulations, all quantities are distributed according to the underlying ensemble. Therefore, the mean of every quantity over all configurations weighted by the number of configurations leads to the correct averages in practice. The excess contributions then become

$$\mu^{ex} = -k_\mathrm{B}T \ln\left(\frac{1}{n_i n_c \langle V\rangle}\sum_{n_i}\sum_{n_c} V\mathrm{e}^{-\beta\Delta U}\right), \tag{9}$$

where n_i and n_c denote the number of insertions per configuration and the number of configurations, respectively. It is important that n_c and n_i are large in order to properly sample the configuration space of the N-particle system as well as the energy landscape of each configuration. In Section 3.1 we will demonstrate this method using the explicit example of a Lennard-Jones fluid.

2.2 Thermodynamic integration

As we will show in Section 3.1, the TPI method works efficiently for not too dense systems when the shape of the particles can be assumed to be spherical. An alternative method for the calculation of chemical potentials, which is efficient also for dense systems, is Thermodynamic Integration (TI)[13]. We start with the definition of the Helmholtz free energy of a system,

$$\begin{aligned} F(N,V,T) &= -k_\mathrm{B}T \ln Q(N,V,T) \\ &= -k_\mathrm{B}T \ln \frac{V^N}{\Lambda^{3N}N!}\prod_{k=1}^{N}\int_0^1 \mathrm{d}^3\mathbf{s}_k \mathrm{e}^{-\beta U(\mathbf{s}_1,\ldots,\mathbf{s}_N;\mathbf{L})}, \end{aligned} \tag{10}$$

where we used the same notation as in Sec. 2.1. How does this free energy change if we vary the interactions in the system? To answer this, we make the potential energy dependent on a coupling parameter λ, i.e. $U(\mathbf{s}_1, \ldots, \mathbf{s}_N; \mathbf{L}) = U(\lambda)$. By tuning λ from 0 to 1, the corresponding interaction can be switched on. Generally, λ can couple to any sort of interaction between any particles in the system, for example it can be used to transform an ideal gas to a Lennard-Jones fluid. Also, the dependence of the energy on this coupling parameter is arbitrary, however due to its simplicity, we assume a linear relation between λ and $U(\lambda)$:

$$U(\lambda) = \lambda U(1) + (1-\lambda)U(0) = U(0) + \lambda(U(1) - U(0))\,. \tag{11}$$

We now evaluate the difference in free energy between the states $U(0)$ and $U(1)$, i.e. when turning on the interactions, and use the integral

$$F(\lambda = 1) - F(\lambda = 0) = \int_0^1 \mathrm{d}\lambda \left(\frac{\partial F(\lambda)}{\partial \lambda}\right)_{N,V,T}. \tag{12}$$

With the help of Eq. (10), the integrand can be calculated as

$$\begin{aligned}
\left(\frac{\partial F(\lambda)}{\partial \lambda}\right)_{N,V,T} &= -k_\mathrm{B}T \frac{\partial \ln Q(N,V,T,\lambda)}{\partial \lambda} = -\frac{k_\mathrm{B}T}{Q(N,V,T,\lambda)} \frac{\partial Q(N,V,T,\lambda)}{\partial \lambda} \\
&= -k_\mathrm{B}T \frac{\prod_{k=1}^{N} \int_0^1 \mathrm{d}\mathbf{s}_k \frac{\partial}{\partial \lambda} \mathrm{e}^{-\beta U(\lambda)}}{\prod_{k=1}^{N} \int_0^1 \mathrm{d}\mathbf{s}_k \, \mathrm{e}^{-\beta U(\lambda)}} = \frac{\prod_{k=1}^{N} \int_0^1 \mathrm{d}\mathbf{s}_k \, \mathrm{e}^{-\beta U(\lambda)} \frac{\partial U(\lambda)}{\partial \lambda}}{\prod_{k=1}^{N} \int_0^1 \mathrm{d}\mathbf{s}_k \, \mathrm{e}^{-\beta U(\lambda)}} \\
&= \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{N,\lambda},
\end{aligned} \tag{13}$$

where the index λ in the subscript of the brackets $\langle ... \rangle_{N,\lambda}$ indicates that the canonical average has to be taken in the system with the potential energy $U(\lambda)$. The latter expression is valid for an arbitrary dependence of U on λ. Using our definition in Eq. (11), the free energy difference simplifies to

$$F(\lambda = 1) - F(\lambda = 0) = \int_0^1 \mathrm{d}\lambda \, \langle U(1) - U(0) \rangle_{N,\lambda}\,. \tag{14}$$

This expression is most frequently used in practice, as it is rather simple to evaluate and most potentials used in computer simulations depend only linearly on interaction constants, such as charges or Lennard-Jones energy parameters.

In order to compute the integral in Eq. (14), we perform a sequence of n_λ simulations, each with a different value λ. In each of these simulations, we sample the average value of the derivative $\langle \partial U / \partial \lambda \rangle$. Finally, we numerically integrate these values over λ. To do this, we have to ensure that the integrands in Eqs. (12) and (14), respectively, are smooth functions of λ and also sampled well. Often, an equidistant choice of the values of λ is sufficient to get a good estimate of the integral. However, if there are intervals on which the integrand is changing considerably, more values of λ within those intervals are necessary to

ensure good sampling. It can be shown that Eqs. (12) and (14) also apply to NpT ensembles by just interchanging the Helmholtz free energy F with the Gibbs free energy G.

The main disadvantage of the TI method over the TPI is the amount of computational time needed to calculate the free energy differences over a sequence of λ values. There are many other different methods to calculate the free energy difference between two states, for example the Exponential averaging method (EXP)[14], the Bennett Acceptance Ratio method (BAR)[15] and the more sophisticated Multistate Bennett Acceptance Ratio method (MBAR)[16], which all are based on sampling the differences in potential energy between two states and calculating the free energy from it. As in the TI, several trajectories are needed for a proper sampling of the free energy differences. For the interested reader we refer to the literature[2,14–16] for more details.

3 Simulation Results: Chemical Potentials in Simple Liquids

3.1 Homogeneous LJ fluid

We will now apply the methodology introduced in the previous sections in order to determine the chemical potential of a simple model system. As such we consider 1728 particles interacting only via a Lennard-Jones potential and take parameters corresponding to liquid argon at a temperature of 75 K. With the help of such simple toy-models, we can study the computational procedure and convergence properties in detail at comparatively low computational expense. The interaction parameters of argon are taken from the GROMOS53A6 force-field[17] and all simulations are performed using version 4.6.7 of the GROMACS simulation package[18,19]. A time-step of 2 fs is used and the velocity-rescale algorithm[20] is applied in order to couple the system to a heat bath. For simulations in the NpT ensemble a Berendsen pressure coupling is employed to rescale the box coordinates accordingly. Since argon is neutral and mono-atomic, in classical MD simulations it is sufficient to model the potential energy via Lennard-Jones (LJ) interactions only,

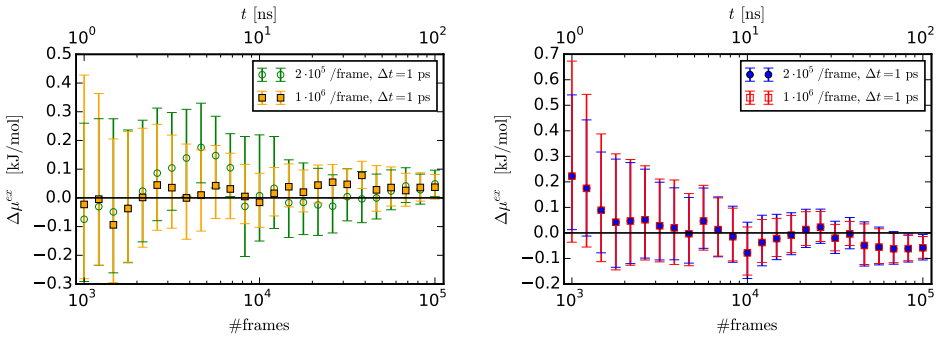
$$U_{LJ}(\mathbf{r}^N) = \sum_{i>j} \frac{C_{12}}{r_{ij}^{12}} - \frac{C_6}{r_{ij}^6}\,, \tag{15}$$

where $C_{12} = 9.696 \times 10^{-6}\,\mathrm{kJ\,mol^{-1}\,nm^{12}}$ and $C_6 = 6.221 \times 10^{-3}\,\mathrm{kJ\,mol^{-1}\,nm^6}$ are the LJ parameters for argon and r_{ij} is the distance between two particles i and j. In MD simulations, the motion of the particles is determined by integrating the forces and only the energies depend on the potentials. Due to computational practicability, those forces are cut beyond a radius r_c. When computing free energies or chemical potentials, we have to ensure that the interaction potential, which is related to the force via integration, is smooth at the cut-off radius[21]. We therefore have to use the spherically truncated and shifted potential, which is shifted by its value at the cut-off,

$$U_{ST}(r) = \begin{cases} U_{LJ}(r) - U_{LJ}(r_c) & r < r_c, \\ 0 & r > r_c. \end{cases} \tag{16}$$

For homogeneous systems, this shift will result in a shift of the chemical potential, thus it may be neglected if one is interested in free energy differences only. However, for inhomogeneous systems it is important to take this potential shift into account.

In our simulations we cut off the Lennard-Jones interactions at $r_c = 0.9\,\mathrm{nm}$ and every $1\,\mathrm{ps}$ all the particle positions are saved for later analyses. Using these saved configurations the test particle insertions are performed later on. In Figure 1 we show the deviation of the excess chemical potential from its reference value, which was determined independently via TI, for 2×10^5 and 1×10^6 insertions in each frame (i.e., each stored configuration), respectively, both for NVT and NpT simulations. Confidence intervals were computed using a Bayesian estimation method[22].



(a) NVT ensemble

(b) NpT ensemble

Figure 1: Running average of the chemical potential of argon in the (a) NVT and (b) NpT ensemble. Shown is the deviation from the reference value $\mu^{ex} = -3.689\,\mathrm{kJ/mol}$ obtained via thermodynamic integration.

For the Lennard-Jones fluid we observe that after 10^5 frames the excess chemical potential has converged to a small statistical uncertainity using 2×10^5 insertions per frame in both ensembles. Using 5 times more insertions in each frame does not yield any significant improvement, thus we conclude that the phase space is sampled sufficiently with about 10^{10} total insertions. However, upon close inspection a slightly better convergence can be observed in the NpT ensemble compared to the NVT ensemble, which we attribute to the more frequent formation of cavities during the rescaling of the coordinates due to the pressure coupling. Note that for spherically symmetric particles, the phase-space is sampled quickly due to the fact that the orientation of the test particle does not play a role. This is different for water, as we will discuss later on.

3.2 LJ fluid at vapor–liquid coexistence

After assessing that chemical potentials in a bulk LJ fluid can be measured with high accuracy using the TPI method, we now turn to an inhomogeneous model system, in which we measure the spatially resolved chemical potential across an interface. This poses a much more stringent test of the methodology, as we know from basic statistical mechanics that the chemical potential is uniform in an equilibrium situation, which, however, involves massive cancellation between ideal and excess contributions. The interface is modeled in a simple way by imposing an inhomogeneous density distribution, for which we employ

the system of liquid argon again, where the fluid phase is now brought into contact with a vapor phase. At the one end of the simulation box we confine the liquid by a wall at $z = 0$ interacting via an integrated Lennard-Jones potential with all particles, which corresponds to a bulk liquid. A second wall is added at $z = L_z$, which is purely repulsive in order to model a vapor phase. In this setup, periodic boundary conditions are applied only in the x and y directions. A snapshot of the system is shown in Fig. 2.

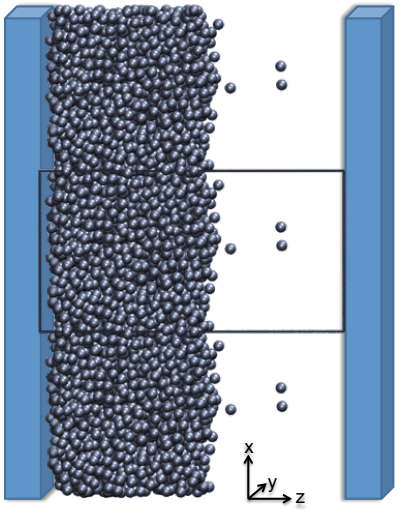


Figure 2: Snapshot of the argon liquid–vapor system consisting of $N = 1728$ particles. The simulation box (solid black line) of shape $L_x = L_y = 3.838\,\mathrm{nm}$, $L_z = 5.549\,\mathrm{nm}$ is periodically replicated in the x and y dimensions. Walls are indicated in blue.

We now want to demonstrate the fact that in thermal equilibrium the chemical potential of any substance is homogeneous in space and constant in time. To this end, we decompose the chemical potential according to Eq. (8) into ideal and excess parts, however, we have to consider that the contributions $\mu^{id} = \mu^{id}(\mathbf{r})$ and $\mu^{ex} = \mu^{ex}(\mathbf{r})$ now depend on the particle density ρ and hence on the position (see Fig. 3). We therefore divide our system in slabs parallel to the x-y plane, with a small thickness Δz, and employ the TPI in these pre-defined slabs separately. For this purpose we use a modified TPI code[23] within the GROMACS software package. Note that at the interface, the density varies quickly, hence the slab width Δz has to be chosen as small as 0.1Å to get a proper resolution of the density profile. The particle density within the individual slabs has been determined using the MDAnalysis package[24].

From the density profile (solid black line in Fig. 3) it can be seen that the interface is located roughly at $z = 4.5\,\mathrm{nm}$ and is about $1\,\mathrm{nm}$ wide. The apparent interfacial width results from a combination of capillary roughness effects and the intrinsic width of the interface[25]. The density in the liquid phase of $\rho \approx 1400\,\mathrm{kg/m^3}$ drops down to $\rho \approx 4\,\mathrm{kg/m^3}$ in the vapor phase (inset in Fig. 3). As shown in Fig. 3, the excess chemical potential

increases from its liquid bulk value of $\mu^{ex} = -3.689\,\mathrm{kJ/mol}$ to a value of about zero in the vapor phase. This means that due to the low density in the vapor, the test particle rarely finds interaction partners and hence the interaction energies are low. On the other hand, one observes the ideal part to drop across the interface, thereby exactly compensating the excess part. The ideal gas contribution follows from Eqs. (6) and (7) as

$$\mu^{id}(\rho) = k_\mathrm{B}T \ln\left(\frac{\Lambda^3(N+1)}{V}\right) \approx k_\mathrm{B}T \ln\left(\rho\Lambda^3\right) = k_\mathrm{B}T \ln\,\rho + k_\mathrm{B}T \ln\,\Lambda^3. \tag{17}$$

The choice of the parameter Λ is completely arbitrary in classical simulations as it yields only an offset in the chemical potential. For our purposes, we choose the thermal wavelength for argon,

$$\Lambda = \frac{h}{\sqrt{3k_\mathrm{B}T}} \approx 0.0461\,\mathrm{nm}. \tag{18}$$

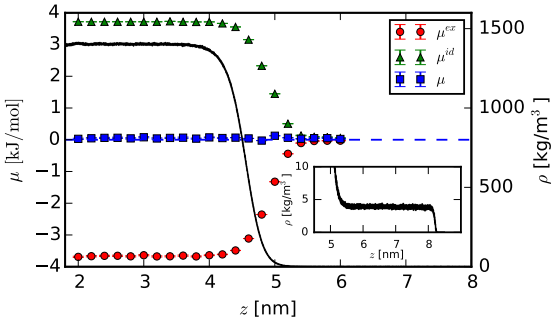


Figure 3: Chemical potential and density profile of the argon liquid–vapor interface. The sum of the excess part (circles) and ideal part (triangles) results in the total chemical potential (squares). The dashed blue line serves as guide to the eye. The solid line represents the density profile for orientation. In the inset the vapor density is shown.

Interestingly, the changes in the excess and ideal contributions to the chemical potential start at the interfacial position and extend approximately $1\,\mathrm{nm}$ into the vapor phase. This long-ranged effect of the excess chemical potential can be associated with the pronounced surface roughness due to interface fluctuations, which depends sensitively on the size of the simulation box[25]. As the shape of the density profile significantly contributes to the chemical potential, great care has to be taken in the correct spatial resolution of the individual contributions. We see that with the used simulation methods, the total chemical potential is within high accuracy constant across the interface, as expected.

3.3 Bulk water

For most biological processes water serves as a solvent, which is typically in equilibrium with a reservoir (for instance the cytosol). Hence, the determination of the water chemical potential in computer simulations is of particular interest. However, measurements

of the chemical potential of molecules in condensed systems such as biological matter is challenging.

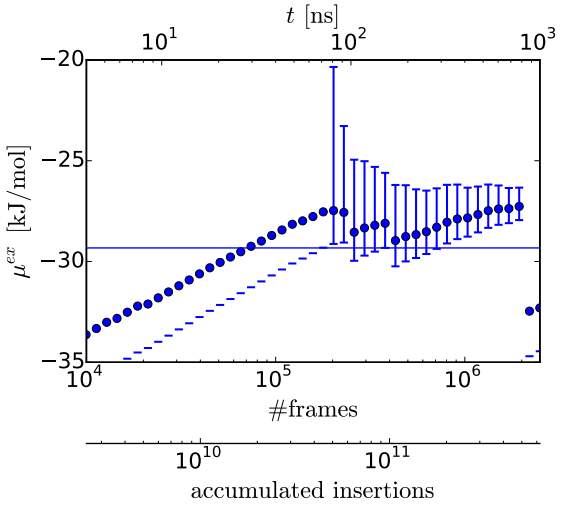
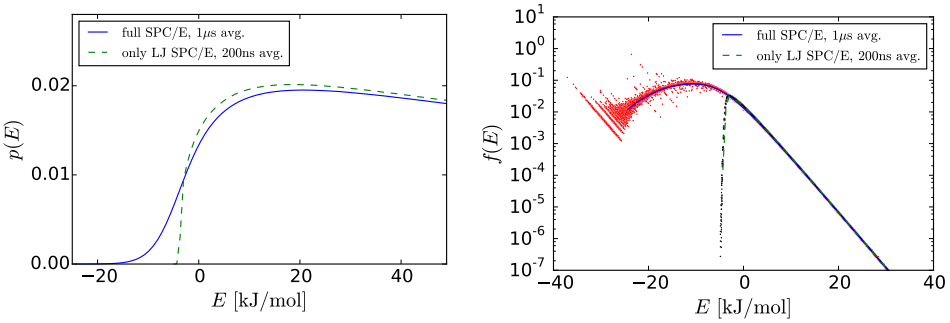


Figure 4: Running average of the excess chemical potential of SPC/E water in the NpT ensemble calculated via TPI. For each frame 2.5×10^5 insertions are performed. Error bars are only drawn when the upper bound is smaller than $-20\,\mathrm{kJ/mol}$, otherwise only the lower bound is indicated. The horizontal line represents the bulk value $\mu^{ex} = -29.33\,\mathrm{kJ/mol}$ determined independently via thermodynamic integration.

Figure 4 shows the result for the water excess chemical potential determined via TPI in a cubic simulation box consisting of 1728 SPC/E[7] water molecules at $T = 300\,\mathrm{K}$ and atmospheric pressure. The simulation was performed for a total time of $1\,\mu\mathrm{s}$ with a write-out frequency $\Delta t = 0.4\,\mathrm{ps}$, on which the insertions are performed. Even for as many as 2.5×10^5 insertions/frame, we observe only poor convergence of the excess chemical potential to its bulk value of $-29.33\,\mathrm{kJ/mol}$ as determined independently via thermodynamic integration (indicated by the horizontal solid line in Fig. 4). The problem with TPI of non-spherical molecules into dense systems is that the number of successful insertions (i.e. insertions with low energies) has to be large enough to sample the phase space sufficiently. In the case of non-spherical molecules, also the rotational degrees of freedom need to be considered and for liquid water at room temperature the probability of finding favourable configurations is correspondingly rather low.

We illustrate these problems of TPI for water by computing the probability distribution $p(E)$ of insertion energies $E = \Delta U$ generated by inserting a single water molecule at random position and with random orientation into a given configuration of water molecules that comes from the simulation trajectory. What one is actually averaging over in Eqs. (6) and (7) is the Boltzmann factor, so the integrand is actually $\mathrm{e}^{-\beta E} p(E) =: f(E)$. While $p(E)$ is somewhat small for negative E, as shown in Fig. 5a, the integrand $f(E)$ in Fig. 5b is not, which generates significant sampling problems. The phase space is sampled sufficiently only if $f(E)$ is sampled well across its maximum value. Otherwise, the computed thermodynamic potential will not be correctly sampled[11,12].



(a) insertion probability distribution

(b) value of $f(E) = \mathrm{e}^{-\beta E} p(E)$

Figure 5: a) Probability distribution p(E) and (b) Boltzmann-weighted probability distribution f(E) for an entire SPC/E water molecule (averaged over a 1 μs trajectory) and for its LJ contribution only (averaged over a 200 ns trajectory). In (b) we also show red and black data points determined from individual trajectories of 10 ns length for the entire water molecule and the LJ sphere only, respectively.

These sampling issues are particular for insertions of an entire water molecule, as can be seen by comparison with the case of inserting only the LJ part of the SPC/E water molecule presented in Figure 5. The distribution of insertion energies in Fig. 5a shows that for an entire water molecule, the probability of finding strongly negative insertion energies is not negligible. These negative insertion energies correspond to insertions of a randomly oriented water molecule in such a way that the interaction with the charges of neighbouring water molecules is favorable. The problem is that the large contribution to the integrand $f(E)$ results from the multiplication of the slowly decaying probability $p(E)$ shown in Fig. 5a with a large number resulting from the Boltzmann weight in $f(E)$ and is thus badly sampled. In fact, the jumps seen in Figure 4 are caused by rare events of finding a suitable cavity with large negative insertion energies. However, if only the neutral LJ sphere of the SPC/E water model is considered for insertions, the probability distribution $p(E)$ quickly decays for negative insertion energies and $f(E)$ has a pronounced maximum, which is well sampled. This explains why the chemical potential of water is much more difficult to estimate with high precision than the chemical potential of spherical solutes.

Application of advanced sampling methods[11,26–28] can yield improved results. Daly et al.[29] recently reported convergence of bulk water TPI for TIP4P water at $470\,\mathrm{K}$ using a 127 ns trajectory and 2.6×10^7 insertions/frame (write-out frequency 1 ps), however the reduced density at this temperature seems to significantly improve convergence. In the case of SPC/E water at 300 K, we do not observe satisfactory results for nearly 10^{12} total insertions, as demonstrated in Figure 4.

The analysis of the interaction between solvated surfaces, which we will discuss below, will require an accuracy of about $0.05\,\mathrm{kJ/mol}$ in the chemical potential (see Section 4). As this will obviously demand much more than 10^{12} insertions in total, we consider it computationally unfeasible to use TPI with whole water molecules. Instead, we determine the excess chemical potential due to Lennard-Jones interactions and due to electrostatic interactions separately. In Fig. 6 we show that using TPI, the LJ contribution to the chemical potential of a water molecule converges quickly to an uncertainity of about $0.02\,\mathrm{kJ/mol}$

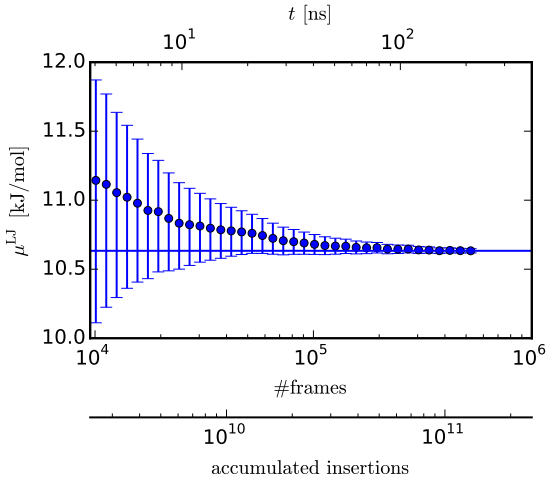


Figure 6: Running average of the Lennard-Jones contribution to the excess chemical potential of bulk SPC/E water, corresponding to to the chemical potential of a neutral LJ particle.

within less than 50 ns. Indeed, Fig. 5a shows that the phase-space for an inserted neutral LJ particle is sampled properly, and the Boltzmann-weighted insertion probability in Fig. 5b exhibits a clear maximum that is well sampled.

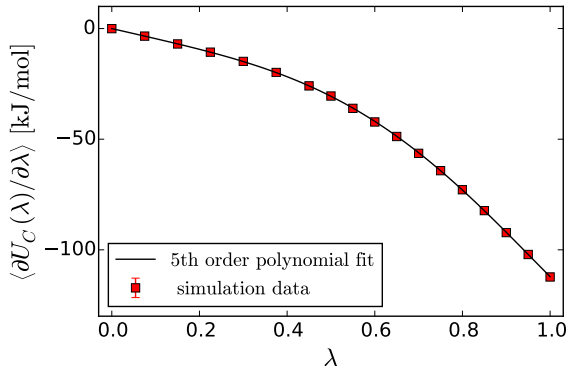


Figure 7: Partial derivative of the electrostatic interaction energy U_C of a SPC/E water molecule with respect to the scaling parameter λ, corresponding to the integrand for thermodynamic integration in Eq. (12). Here, $\lambda = 0$ corresponds to a molecule without partial charges and at $\lambda = 1$ all charges are present. The symbols represent simulation results at discrete values of λ and the solid line is a 5th order polynomial fit to the data points. Simulations are performed in a box consisting of 1728 water molecules in the NpT ensemble.

In order to determine the electrostatic contribution, we employ a thermodynamic integration procedure, in which only the electrostatic interactions are changed. This is advantageous as the LJ interactions can be determined independently with high accuracy

using the computationally much cheaper particle insertion method. For our simulations we use a non-uniform sampling for the λ-values, where λ is a prefactor that uniformly scales the partial charges of water from zero to the corresponding values, which we found to be numerically efficient and to yield high accuracy (see Fig. 7). The integration can be performed in two distinct ways, fitting a fifth-order polynomial followed by analytic integration or application of Simpson's rule for the integration. The influence of the integration method on the final result is negligible, as is shown in Fig. 8.

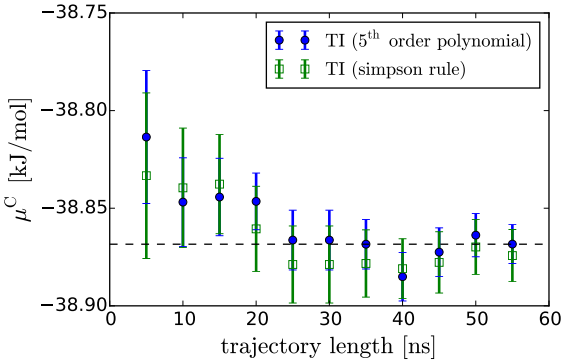


Figure 8: Electrostatic contribution to the chemical potential of SPC/E bulk water in dependence of the simulation length. Simulations are performed in the NpT ensemble using a simulation box containing 1728 water molecules. The dashed line serves as a guide to the eye.

In Fig. 8 the electrostatic contribution to the chemical potential of bulk water is shown. Convergence to an uncertainity of $0.01\,\mathrm{kJ/mol}$ is obtained within 50 ns simulation time for each of the simulations at distinct λ-values. We can now estimate the total simulation time required to determine the chemical potential of water within the desired accuracy of $0.05\,\mathrm{kJ/mol}$: 18 simulations at distinct λ ($\sim 50\,\mathrm{ns}$ each) for the TI, a trajectory of length $\sim 50\,\mathrm{ns}$ for the TPI, plus the actual TPI analysis, which roughly requires the same computational effort as needed for generating the actual trajectories, thus an equivalent of another $\sim 50\,\mathrm{ns}$. We therefore obtain an estimate for the total simulation time of the order $20 \times 50\,\mathrm{ns} = 1000\,\mathrm{ns}$ for the total computational costs of determining the chemical potential of water with this high accuracy.

Splitting the excess chemical potential into its electrostatic and Lennard-Jones contributions,

$$\mu^{ex} = \mu^{\mathrm{LJ}} + \mu^{\mathrm{C}}, \tag{19}$$

has proven to be a reliable method for efficiently determining the excess chemical potential of water with high accuracy and making optimal use of computational resources. As a rule of thumb, on a typical computer node with 4 to 8 physical cores, e.g. for typical membrane systems hydrated with water, consisting of 72 atomically resolved lipid molecules and about 1500 water molecules, 10–50 nanoseconds a day can be achieved. Such simulations are still expensive, but computationally feasible.

4 Surface Interactions at Prescribed Water Chemical Potential

4.1 The interaction of surfaces in water

In the last chapter we have presented the methodology to measure the chemical potential of water in computer simulations with high accuracy. We will now turn to an example where the precise control of the chemical potential is of crucial importance, namely the interaction of extended planar surfaces in aqueous solution. Biological tissues and cells are composed of diverse functional units such as organelles, protein complexes, and carbohydrate assemblies. The structural organization of these cellular constituents on the sub-micrometer scale is essential for their proper function and in the congested biological environment largely depends on the physical interactions between their surfaces. Without regulation of these interactions by the organism, essential cellular processes such as material transport or cell division would not be possible. In an analogous way the behaviour of colloidal solutions in technology and micro-emulsions are governed by interactions between surfaces with different molecular compositions. In the following we will first introduce the concept of surface interactions, and then discuss a method to determine the interaction pressure at prescribed water chemical potential in MD simulations. After describing our computer models of interacting surfaces we will present some background on the fundamental surface interactions observed in these systems, namely vdW and hydration forces. Finally, a detailed discussion of the simulation results will be given and we show that the simulation protocol enables us to draw conclusions on the interaction mechanisms. At the end of this section we reproduce the experimentally observed interaction characteristics of phospholipid membranes and shed light on underlying mechanisms of the ubiquitous hydration repulsion between biological and technical surfaces.

The interaction between planar extended surfaces in solution is commonly described in terms of pressure–distance relations, where the interaction pressure Π is measured as a function of the surface separation D. Π is in fact related to the derivative of the Gibbs free energy, G, with respect to D. When the surface area A is kept constant and D is varied, while the system is in contact with a bulk water reservoir of temperature T and water chemical potential μ, the corresponding pressure–distance relation reads:

$$\Pi_{\mu,T,A}(D) = -\frac{1}{A}\left(\frac{\partial G}{\partial D}\right)_{\mu,T,A}. \tag{20}$$

Importantly, μ is dictated by the reservoir and is independent of D in this scenario. The interaction pressure $\Pi_{\mu,T,A}(D)$ acting between two surfaces sensitively depends on the chemical potential of the confined solvent, as it determines the number of molecules present between the surfaces. Accordingly, if there is an excess of particles, the pressure acquires a repulsive component, whereas for a lack of particles an attractive contribution is generated. Simulating two interacting surfaces at prescribed water chemical potential is generally a challenging task. Explicitly simulating the bulk reservoir is computationally very demanding and does typically have a number of disadvantages[30]. An alternative approach consists of Grand-Canonical Monte Carlo (GCMC) simulations, where the chemical potential is continuously adjusted through stochastic deletion and insertion of water molecules[31]. However, for hydrophilic surfaces exhibiting very dense water layers, the stochastic deletion/insertion algorithms in GCMC generally deliver poor convergence,

usually with a moderate pressure resolution of about 100 bars[32]. The reason for this is similar to the reason why TPI exhibits sampling problems, see our discussion in Section 3.3. In the next Section we discuss an alternative computational method that extrapolates the interaction pressure at prescribed chemical potential and that allows for higher pressure resolution.

4.2 Thermodynamic extrapolation

One way of resolving the problem of determining interaction pressures at prescribed chemical potential is extrapolation from a given pressure p at an arbitrary but known chemical potential μ, which we refer to as thermodynamic extrapolation method (TEM)[33]. Let us consider an MD simulation performed at constant pressure p, temperature T and a fixed number of water molecules N_w. In general, the water chemical potential $\mu(N_w, p, T)$ in such a simulation will deviate from the bulk chemical potential μ_0. The total differential of the Gibbs free energy reads as

$$\mathrm{d}G(N_w, p, T) = -S\mathrm{d}T + V\mathrm{d}p + \mu\mathrm{d}N_w, \tag{21}$$

from which we conclude that

$$\mu(p, N_w, T) = \left.\frac{\partial G}{\partial N_w}\right|_{p,T}. \tag{22}$$

As the temperature is kept constant in all of our simulations, we will drop the temperature dependence of all quantities in the following. To investigate the change of the water chemical potential due to a change of the pressure, we now expand the chemical potential in a Taylor series around the pressure p_0 for constant N_w,

$$\Delta\mu(p, N_w) = \left.\left(\frac{\partial\mu}{\partial p}\right)_{N_w}\right|_{p=p_0} \Delta p + \frac{1}{2}\left.\left(\frac{\partial^2\mu}{\partial p^2}\right)_{N_w}\right|_{p=p_0} (\Delta p)^2 + \dots . \tag{23}$$

From the total differential of Eq. (21) we obtain the Maxwell relation

$$\left(\frac{\partial\mu}{\partial p}\right)_{N_w} = \left(\frac{\partial V}{\partial N_w}\right)_p = v_w(p, N_w)\,, \tag{24}$$

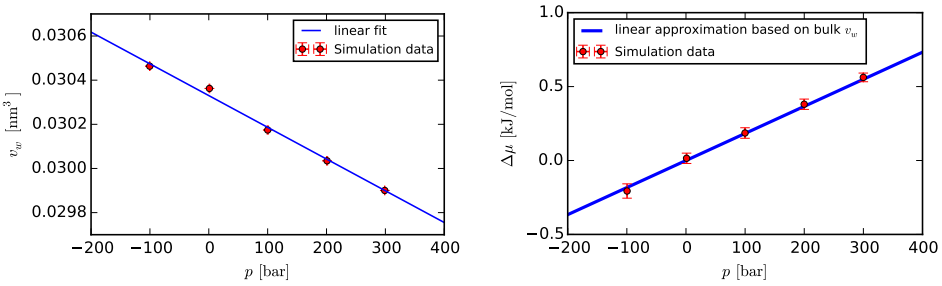
which we use to define the partial volume v_w of a water molecule. Note that in heterogeneous confined systems, water molecules interacting with other species in general can exhibit a different partial volume compared with the bulk value. In Fig. 9a we show the partial volume of bulk water at various pressures and find $v_w(p = 1\,\mathrm{bar}) = 0.0304\,\mathrm{nm}^3$. Insertion of Eq. (24) into Eq. (23) yields

$$\Delta\mu(p, N_w) = v_w(p, N_w)\Delta p - \frac{1}{2}v_w(p, N_w)\kappa(p, N_w)\,(\Delta p)^2 + \dots\,, \tag{25}$$

where in the quadratic term we intrduced the compressibility of water,

$$\kappa(p, N_w) = -\frac{1}{v_w(p, N_w)}\left.\frac{\partial v_w(p, N_w)}{\partial p}\right|_{N_w}. \tag{26}$$

Using $\partial v_w/\partial p = -1.437 \times 10^{-6}\,\mathrm{nm}^3/\mathrm{bar}$, as determined from a linear fit in Fig. 9a, we obtain the compressibility of the SPC/E water model at room temperature and atmospheric



(a) partial water volume

(b) change of the chemical potential at different pressures

Figure 9: Left: Volume per water molecule v_w as a function of the pressure in bulk water simulations (symbols) and a linear fit (solid line). At atmospheric pressure one finds $v_w = 0.0304\,\mathrm{nm}^3$. Right: Chemical potential of bulk water as a function of pressure. Here, the deviation $\Delta\mu = \mu - \mu_0$ of the chemical potential μ from $\mu_0(p_0 = 1\,\mathrm{bar})$ is plotted. The slope of the solid line is given by the linear term of Eq. (25), $\Delta\mu = 0.0304\,\mathrm{nm}^3\,p$.

pressure from our simulations, $\kappa(1\,\mathrm{bar}) = 4.74 \times 10^{-5}\,\mathrm{bar}^{-1}$, which is in nice agreement with other simulation results[34,35] and the experimental value $\kappa^\star = 4.52 \times 10^{-5}\,\mathrm{bar}^{-1}$ for water at 25°C[36]. Note that due to this low compressibility, the quadratic term becomes comparable to the linear one only at very elevated pressures of more than $40\,\mathrm{kbar}$. We therefore neglect this correction term and remain with the linear term only in the following.

We next use the Maxwell relation, Eq. (24), to determine the partial volume of water via the chemical potential from our simulation data. To this end, we measure the chemical potential μ of water as a function of the applied pressure p. As seen in Figure 9b, the data indeed follows the linear relation as predicted by Eq. (23), with the proportionality factor $v_w = 0.0304\,\mathrm{nm}^3$ being the partial volume of water.

On the basis of Eq. (25) we can now extrapolate the interaction pressure p that we obtained from the simulation at a chemical potential μ to the value p_0 that corresponds to the bulk chemical potential μ_0 (or any other chemical potential we are aiming for). Transforming Eq. (23), we arrive to linear order at

$$p_0 \approx p - \frac{\mu - \mu_0}{v_w}\,. \tag{27}$$

The latter expression thus allows for a comparison between experimental pressure distance curves and those obtained in the simulations in a wide pressure and separation range.

4.3 Computer models of interacting synthetic surfaces

In order to investigate the influence of surface properties on surface-surface interactions across water we use simple yet versatile model surfaces based on self-assembled monolayers (SAMs) of amphiphiles. These molecules consist of $(CH_2)_9\,CH_3$ alkyl chains, terminated by polar hydroxyl (-OH) headgroups. Surface parameters such as the monolayer stiffness and the polarity of the headgroups can be adjusted in order to model, for

instance, rigid or soft, and hydrophilic or hydrophobic surfaces. A snapshot of the simulation architecture is shown in Fig. 10, where 2×100 surface molecules are arranged on a hexagonal lattice at the prescribed area per chain of $0.234\,\mathrm{nm}^2$[37]. The force-field parameters for the simulations are taken from ref.[37] and the SPC/E water model[7] is employed. With this setup, we perform simulations in the NVT ensemble and couple a heat bath via the velocity-rescale algorithm[20]. For long-range electrostatic interactions, the PME method[38,39] is applied and short range interactions are cut-off at 0.9 nm.

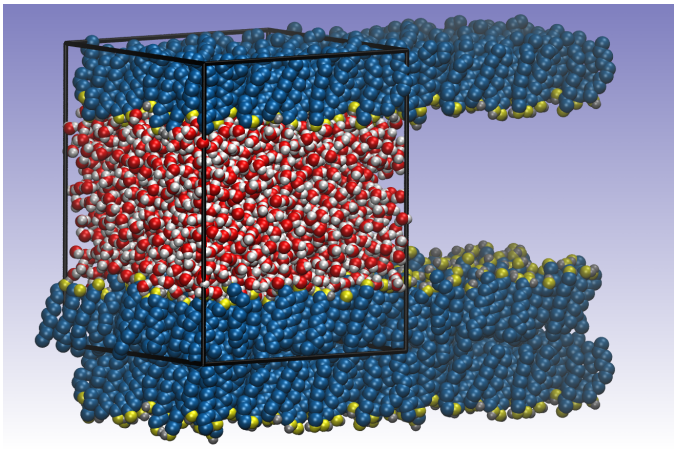


Figure 10: Simulation snapshot of self-assembled monolayers of defined surface characteristics interacting across water. Hydrophobic hydrocarbon chains are shown in blue, the oxygen of the polar headgroups in yellow. Water is shown inside one simulation box only for clarity.

The separation between opposing surfaces, D, is varied by adjusting the box dimension L_z and the number of waters, N_w correspondingly. As we have seen in Section 3.2, the excess and ideal contributions of the chemical potential are not constant in space, but only their sum. We therefore always determine the chemical potential by measurements of μ^{id} and μ^{ex} at the same distance from the surfaces, which for practical convenience is chosen as the center of the water layer.

For the thermodynamic integration, which is used to determine the Coulomb contribution to the chemical potential, we restrain the position of the integrated particle to this position by applying a strong harmonic potential. We verified that this does not lead to any artifact by comparing to simulation results where no restraints are applied to the coordinate, but their position is kept fixed. The procedure exactly reproduces the results when restraints are applied, thus all interactions are accounted for correctly. Note however, that in constant pressure simulations such fixed coordinates can lead to various artifacts. For the TPI of the LJ spheres, the insertions into predefined slabs are performed as already described above.

4.4 Interaction mechanisms

Before presenting simulation results, we quickly remind the reader of the most basic interaction mechanisms between neutral solvated surfaces, namely van der Waals (vdW) and hydration forces. They are sufficient to describe interactions between the charge-neutral SAM surfaces introduced above and also between phospholipid membranes as discussed further below. Another interaction that gives rise to large contributions on nanometer lengthscales are electrostatic effects if the surfaces are not charge-neutral or if salt ions are present in the solvent mixture. Note however, that the interaction between molecularly complex biological surfaces such as biomembranes is typically the result of a subtle interplay involving also other contributions of comparable magnitude, which cannot be discussed within this chapter[40].

4.4.1 Van der Waals interactions

The electric charge distribution (due to nuclei and electrons) of atoms and molecules manifests in the form of an interaction that acts among all matter, even when the constituting parts are charge-neutral. A simplified picture, which has been established in the 1930s, attributes this interaction to three kinds of atomic dipole–dipole interactions[41]:

- interactions between two permanent molecular dipoles (Keesom interactions),
- interactions between a permanent molecular dipole and a corresponding induced molecular dipole (Debye interactions),
- interactions between two spontaneously induced molecular dipoles (London dispersion interactions).

The unified concept of van der Waals (vdW) interactions accounts for all three cases, and in each case the interaction potential, i.e., the work needed to bring two molecules from infinite separation to a certain distance r, varies as the inverse-sixth power of distance,

$$u(r) = -\frac{C}{r^6}, \tag{28}$$

where C is a positive coefficient. Biomembranes are extensive bodies composed of atoms and molecules, which all interact via such vdW potentials. In a first approximation, we can assume pair-wise additivity of the individual interactions and sum them all up. In the case of two parallel semi-infinite media separated by distance D in vacuum, we obtain the free energy per unit area,

$$F_{\mathrm{vdW}}/A = -\frac{H}{12\pi D^2}, \tag{29}$$

where $H = \pi^2 C n^2$ is known as the *Hamaker constant*[42] and n corresponds to the particle number density of the two media. In the case when the media are immersed into water, additional contributions appear,

$$H = \pi^2 C n^2 - 2\pi^2 C_{\mathrm{cw}} n n_{\mathrm{w}} + \pi^2 C_{\mathrm{w}} n_{\mathrm{w}}^2, \tag{30}$$

with n_{w} being the number density of water molecules, C_{cw} the interaction coefficient between water and the particles constituting the semiinfinite media, and C_{w} the coefficient

for the water–water interactions. The vdW pressure between the membranes can then be obtained by derivation of the free energy, $p_{\mathrm{vdW}} = -\partial(F_{\mathrm{vdW}}/A)/\partial D$, which yields

$$p_{\mathrm{vdW}} = -\frac{H}{6\pi D^3}. \tag{31}$$

The resulting vdW pressure decreases with the inverse cube of the separation. Van der Waals forces between two bodies of the same material are always attractive, but are repulsive when two different materials interact across a third material with intermediate dielectric properties[41].

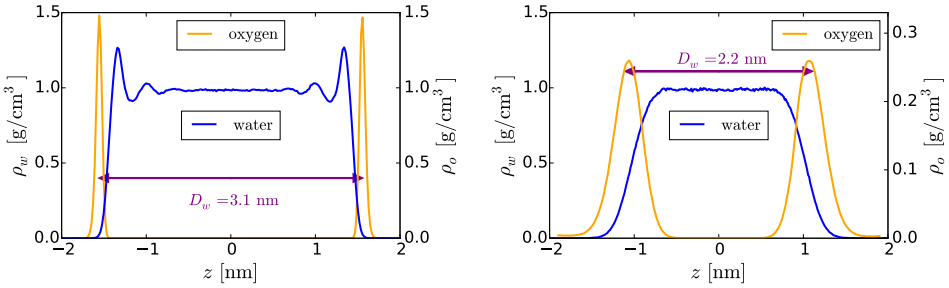
4.4.2 Hydration interactions

When surfaces approach to very short separations that are comparable to the size of a few water molecules, the molecular structures of the surface and of the water layers become important. The removal of the last few layers of water molecules gives rise to a very strong force, the *hydration repulsion*. This interaction was quantified experimentally for lipid membranes in the 1970s in diffraction experiments of membrane multilayers[43,44] and by surface force apparatus measurements of a single pair of membranes[45]. Experiments revealed that the hydration repulsion is a short-ranged force, acting typically on the nanometer range. It approximately obeys an exponential decay law with a decay constant λ between 0.1 and 0.5 nm[46]. The hydration repulsion can easily reach pressures of several kilobars at small separations, overcoming all attractive van der Waals and electrostatic forces. It is responsible for the stabilization of membrane surfaces and stops vesicles from close contact with cell membranes, making spontaneous fusion difficult[47]. The mechanism of the hydration repulsion has been intensively debated. The first attempt to rationalize the repulsion based on water structuring has been proposed by Marčelja et al.[48]. They predicted an exponential decay of the repulsion based on a coarse-grained order-parameter description, accounting for the structural properties of water confined between the membrane surfaces, in qualitative agreement with computer simulations[49]. Alternatively, the repulsion due to the strong enthalpic adsorption of solvent molecules at the surface has been proposed as a possible mechanism for hydration interactions[40]. Finally, the configurational entropy of lipid molecules has been suggested to be responsible for the observed short-range repulsion[50]. Despite the achievements of continuum models, it has become a consensus view during the last decades that in order to explain the hydration interaction on a quantitative level, the water molecules in the hydration layers have to be taken into account explicitly[46]. In recent years, computer simulations have led to a clearer picture about the physical mechanisms responsible for the repulsion. In fact, several of the mechanisms originally proposed were found to contribute[30,33,51,49,40].

4.5 Simulation results for hydrophilic model surfaces

We will now use TEM in order to evaluate the pressure–distance relation for hydrophilic model surfaces. We vary the surface stiffness by restraining individual atoms of the alkane chains by harmonic potentials and scale the headgroup partial charges by a factor α to tune the surface polarity[37].

Figure 11 shows the density profiles of water and headgroup oxygens. For the stiff surfaces in Fig. 11a, the distribution of the headgroup position is narrow and the water density



(a) stiff surface ($\alpha = 0.7$) (b) soft surface ($\alpha = 1$)

Figure 11: Density profiles for the headgroup oxygens (right scale) and water (left scale) between (a) two stiff surfaces with slightly reduced polarity ($\alpha = 0.7$) and (b) two fully polar (α =1) soft surfaces.

profile shows oscillations close to the surface due to water layering. Since the soft surfaces are not entirely rigid, water and headgroup distributions overlap and the headgroups fluctuate by around 0.18 nm around their mean positions (Figure 11b). In both cases we define the surface–surface distance D based on the positions of the density peaks of the headgroup oxygens.

In order to demonstrate the importance of the correct ensemble, we show the extrapolated interaction pressure at fixed chemical potential that corresponds to the value of the bulk reservoir ($\mu = \mu_0$) and compare it to the heuristic assumption that the number of water molecules scales linearly with the headgroup separation, corresponding to the assumption of a constant density of water between the surfaces ($\rho = \text{const.}$),

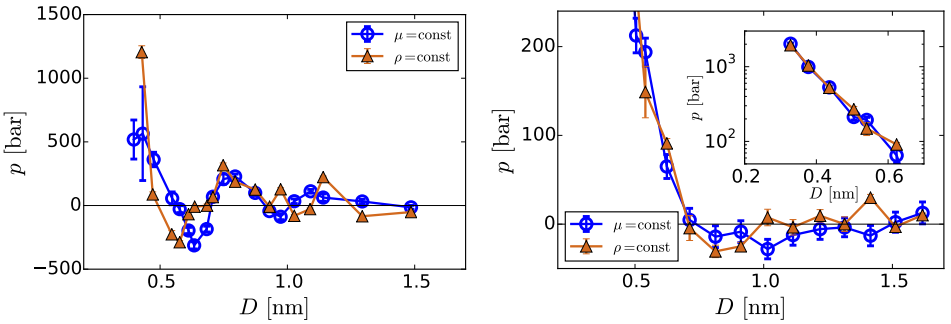
$$N_w(D) = N_w(D_\infty) - \frac{A(D_\infty - D)}{v_w(D_\infty)}. \tag{32}$$

For D_∞ we choose the data point at the largest separation we considered in the simulations, where $p(D_\infty)|_{\mu=\mu_0} = 0$ bar holds within numerical errors.

In Fig. 12 we show the pressure–distance curves for stiff surfaces, which mimic frozen structures with slightly reduced polarities ($\alpha = 0.7$), and a very soft hydrophilic SAMs with full polarity ($\alpha = 1$).

The stiff surfaces (Fig. 12a) exhibit profound pressure oscillations, which are connected to the density fluctuations shown in Fig. 11a. Each oscillation in the pressure–distance profile corresponds to expulsion of exactly one water layer from the interlamellar region[37]. Due to this water layering, the assumption of a constant density between the surfaces in Eq. (32) leads to significant discrepancies in the pressure profile for the stiff surfaces (Fig. 12a) compared with the interactions at fixed chemical potential. In the case of soft surfaces, the pressure (blue curve) does not show oscillatory behavior, but decays exponentially for small separations with a decay length of roughly 0.82 Å. The hypothetical pressure at constant density (green curve) agrees in this case much better with the interaction pressure at constant chemical potential.

In order to further quantify the interactions, we evaluate the free energy profiles by



(a) stiff surface ($\alpha = 0.7$)

(b) soft surface ($\alpha = 1$)

Figure 12: Interaction pressure profiles $p(D)$ between (a) stiff and (b) soft surfaces. In the case of stiff surfaces, the polarity has been slightly reduced by a factor $\alpha = 0.7$. Shown are the interaction pressures at fixed bulk chemical potential as well as a simple model, where the water density is assumed to be constant between the headgroups (Eq. (32)). Lines serve as guides to the eye.

integrating the interaction pressure,

$$F(D) = \int_D^\infty p(D')\,\mathrm{d}D', \tag{33}$$

which accounts only for the work of water removal from the interlamellar region. However, the entire work it takes to bring the surfaces to a certain separations, takes into account also the compression of the SAMs, hence

$$F^{\mathrm{tot}}(D) = \int_D^\infty p(D')\frac{\mathrm{d}L_z(D')}{\mathrm{d}D'}\,\mathrm{d}D'. \tag{34}$$

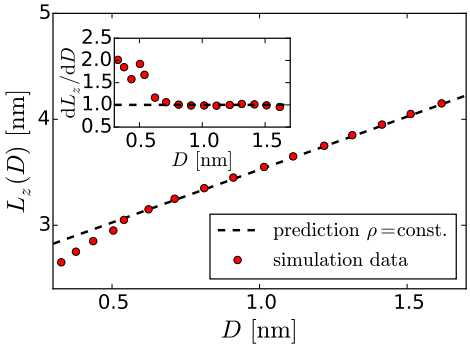
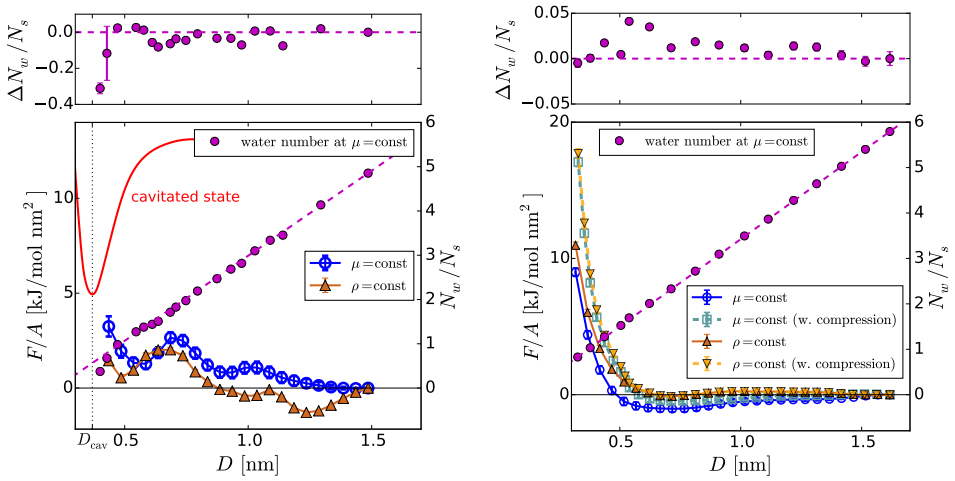


Figure 13: Box length L_z as a function of the headgroup separation D. The dashed line corresponds to the prediction of a constant water density between the headgroups, $L_z(D) = D + L_0$, with L_0 being the length of the simulation box when there is no water present. The inset shows the derivative appearing in Eq. (34).

In the case of stiff surfaces, the compression work is negligible (since $dL_z/dD \approx 1$, not shown). However, for soft surfaces the box length is not a linear function of the headgroup separation at small distances (Fig. 13) and thus the derivative significantly contributes (inset in Fig. 13). Figure 14 shows the corresponding free energies for the stiff and the soft surfaces. For the soft surfaces, quantitative differences appear between Eqs. (33) and (34), see Figure 14b. The main effect shows up only at the smallest separation, below 0.5 nm, where the membrane compression significantly contributes to the entire free energy, as expected from the derivative shown in the inset in Fig. 13, such that the total free energy is higher.

Comparing the actual free energies (blue curves) to the hypothetical free energies based on the constant-density assumption (brown curves in Fig. 14) also shows profound deviations. This illustrates that the SAM system cannot be simply modeled by imposing constant water density. Rather, the exact value of chemical potential needs to be considered.



(a) stiff surface ($\alpha = 0.7$)

(b) soft surface ($\alpha = 1$)

Figure 14: Free energy as a function of distance for the stiff (a) and the soft (b) hydrophilic surfaces at bulk chemical potential (blue spheres). Additionally, the brown triangles show the hypothetical free energy under the assumption of constant water density in the interlamellar region (Eq. (32)). In (a), also the cavitation free energy $F_{\text{cav}}(D)$ is shown (solid red line). For the soft surfaces (b), we compare the free energy of hydration and the total free energy, which includes SAM compression (Eq. (34)) shown as cyan squares and orange triangles for $\mu = \mu_0$ and $\rho = \text{const.}$, respectively. Lines serve as guides to the eye. The right scale shows the corresponding number of water molecules per surface group, N_w/N_s (magenta circles) and the prediction under the constant density assumption (magenta dashed line). On top of both graphs the difference of the water number to the assumption of $\rho = \text{const.}$ is shown for the stiff and the soft surfaces, respectively.

In Figure 14, we also report the number of water molecules per surface group in the simulations at $\mu = \mu_0$. Clearly, the water number (circles) for the stiff surfaces is not bulk-like, as indicated by the dashed line in Fig. 14a which is also represented by the fact that the free energy landscape differs. For the soft surface (Fig. 14b), the predicted water number nearly seems to match, however upon close inspection small differences in

the water number can be observed (top of Fig. 14b), causing the interaction energies to be slightly different, which furthermore corroborates the notion that the correct chemical potential needs to be taken into account in simulations.

In the next step, we consider that the hydrated state might thermodynamically not be stable, i.e. the water could be expelled from the slab. Such a behavior is experimentally known for non-polar and weakly polar surfaces, where polar surface groups are mixed with non-polar ones, and has recently also been analyzed in MD simulations[37]. In order to check the thermodynamic stability of the hydrated state with respect to the cavitated state, we need to compare both respective free energies. The free energy of the cavitated state relative to the hydrated state is evaluated by integrating the interaction pressure p_C of the surfaces across vacuum as

$$F_{\text{cav}}(D) = 2\,(\gamma_{sv} - \gamma_{sw}) + \int_D^\infty p_{\text{cav}}(D')\,\mathrm{d}D', \tag{35}$$

where the first term represents the work to remove water from a (thick) interlamellar region into the bulk, corresponding to the difference in the solid–vapor surface tension γ_{sv} and the surface–water surface tension γ_{sw}. The resulting cavitated free energy is larger than the hydrated free energy (red curve in Figure 14a), therefore the hydrated systems are thermodynamically stable.

We now turn our attention to a more detailed analysis of the interaction of the soft surfaces. In that case, the density profile shows no oscillations and the pressure decays monotonically with distance at small separations and reaches thousands of bars close to contact. However, in the simulations a minimum in the pressure at $D \approx 0.9\,\text{nm}$ can be observed (Figure 15). During the following analysis we will consider only the data where μ = const. Using Eq. (30), one can estimate the static limit of the Hamaker constant for the SAMs assuming that the system can be represented by two semi-infinite halfspaces. The number density of the hydrophobic chains is $n = 38\,\text{nm}^{-3}$ and that of water is $n_w = 33\,\text{nm}^{-3}$. The interaction parameters for the attractive contribution of the Lennard-Jones potential as defined in Eq. (15) are $C_6^{\text{w}} = 2.617 \times 10^{-3}\,\text{kJ mol}^{-1}\,\text{nm}^6$ for water and $C_6^{\text{cw}} = 4.442 \times 10^{-3}\,\text{kJ mol}^{-1}\,\text{nm}^6$ for the chain–water interactions, respectively.

By inserting these numbers into Eq. (30), a simple estimate for the Hamaker constant is $H \approx 5.6 k_{\text{B}}T$. This values agree roughly with literature, where $H \approx 1-10\,k_{\text{B}}T$ is reported for most of the materials[41]. However, for the van der Waals interaction the definition of the interface position is crucial. Reasonable results can be obtained by defining an additional lengthscale d_0 based on the separation of the oxygen atoms on the opposing surfaces. The offset is then determined such that close-contact (without water in between) corresponds to[49] $D - d_0 = 0$. We now compare the predictions of van der Waals and hydration interactions with the simulations in Figure 15. The dash-dotted line represents the vdW attraction according to equation (31) and correspondingly diverges as $D \to d_0$. Thus, for a study of effects at smallest separations also the repulsive contribution of the LJ interaction would need to be considered. Alternatively, here we limit the analysis to distances $D \geq 0.5\,\text{nm}$.

Using $d_0 = 0.26\,\text{nm}$ as determined from independent simulations at $p = 1\,\text{bar}$ of the surfaces interacting accross vacuum[49] and fitting an additional exponential function to the simulation data for $D > 0.5\,\text{nm}$,

$$p(D) = A\,\mathrm{e}^{-D/\lambda} - \frac{H}{6\pi(D-d_0)^3}, \tag{36}$$

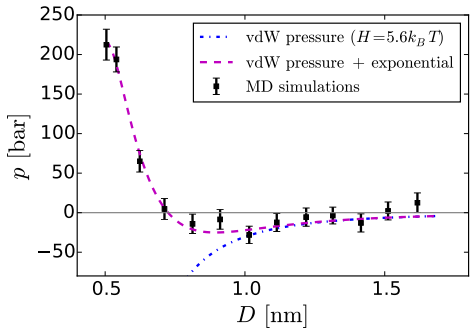


Figure 15: Pressure–distance curve for the soft SAMs. Simulation data are shown as red squares. The dash–dotted blue line represents an estimate for the vdW interaction, the dashed purple line is a combination of the estimated vdW attraction and an exponential repulsion.

we obtain a decay length of $\lambda = 1.0$Å. When considering only the short-range interactions (inset in Fig. 12b) and assuming an exponential decay, the corresponding decay length follows as $\lambda = 0.8$ Å. This nice agreement shows that with the presented methodology of TEM we are able to resolve the attractive contribution of vdW interactions between surfaces as well as the hydration interactions on smaller lengthscales.

4.6 Application to membrane interactions

As boundaries of cells and organelles, membranes play key roles in almost all biological processes: exchange of substances via membrane proteins, interactions between cells, control over the shape of cells, to name just a few. All biomembranes are primarily composed of amphiphilic molecues, mostly lipids, which spontaneously form bilayers in water due to their amphiphilic chemical structure[52] via the hydrophobic effect[53]: The polar headgroups are in contact with water, while the non-polar hydrocarbon chains face towards the bilayer interior, see Fig. 16 for a schematic illustration. The interaction characteristics between biomembranes play important roles in biological processes involving membranes[54]. Predominantly repulsive membrane-membrane interactions, for instance, prevent close membrane-membrane contacts and generate aqueous pathways for the diffusion of biomolecules between the membrane surfaces. Attractive interaction contributions, on the other hand, can create well-defined average membrane separations that facilitate membrane fusion or cell adhesion processes via specific binding sites.

The simplest and most commonly used experimental models of biomembranes are phospholipid membranes. As phospholipids are zwitterionic, they carry larger dipole moments on their headgroups, as opposed to the decanol SAMs analyzed previously. The relatively large dipole moment gives rise to a large hydration repulsion force at small separations, which can reach up to thousands of bars[43,44]. Since the hydration repulsion eludes continuum theoretical treatment, its quantitative description requires rigorous modeling of all molecular components. With the computational techniques presented here, it is now possible to evaluate and study these forces within computer models[33,51]. Figure 16 shows an atomistic computer model of interacting, hydrated phospholipid bilayers composed of

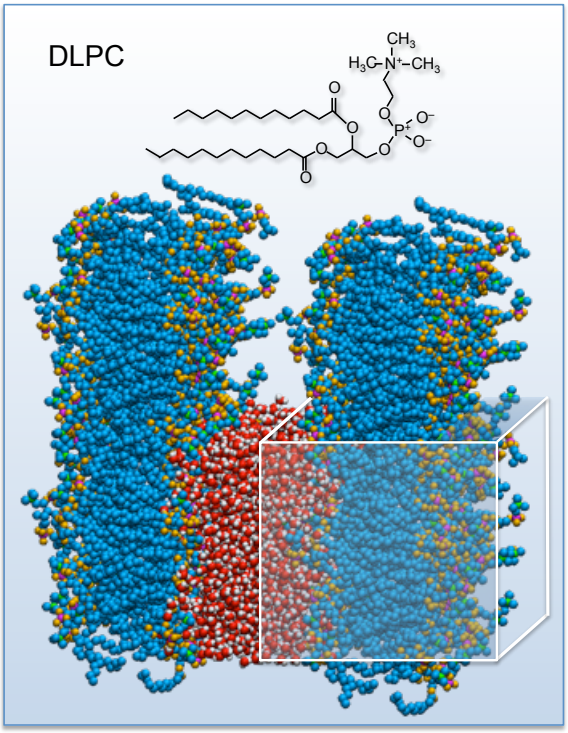


Figure 16: Simulation snapshot of a phospholipid membrane: DLPC bilayer consisting of 72 lipids in the simulation box. Water is only partially shown in the middle section for clarity.

72 dilinoleoylphosphatidyl–choline (DLPC) lipids. The simulations were performed using the velocity-rescale algorithm[20] with a time constant of 1 ps in order to control the temperature. The simulation pressure is controlled by a Berendsen barostat with a time constant of 1 ps and a compressibility $\kappa = 4.5 \times 10^{-5}\,\mathrm{bar}^{-1}$. A Lennard-Jones cutoff at $0.9\,\mathrm{nm}$ is applied and electrostatic interactions are treated using the Particle Mesh Ewald (PME)[38,39] method with a 0.9 nm real-space cutoff. We use the SPC/E water model[7] and the united-atom Berger force field[55].

Fig. 17 shows experimental pressure distance curves of DLPC and DPPC (dipalmitoylphosphatidylcholine) membrane bilayers in comparison with the results of our computer simulations obtained via thermodynamic extrapolation. The simulation results are in remarkable agreement with the experiments. They reproduce the exponential decay law with the experimentally observed decay length, so that we may assume that all dominant repulsion mechanisms are captured by the atomistic computer model. The detailed analysis obtained from computer simulations reveals that the repulsion arises due to several mechanisms, the relative weight of which depends on the surface separation. The results indicate that the long-debated hydration repulsion is not as universal as initially assumed, but rather consists of a mix of competing contributions that change role with varying separation[33,51,49,40].

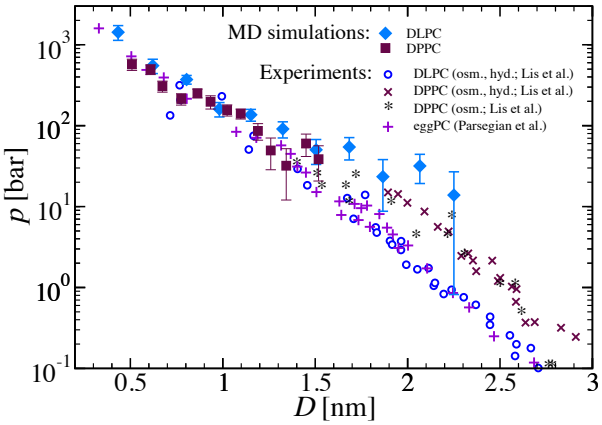


Figure 17: Pressure–distance data for different phospholipids. The obtained simulation data agrees well with experiments[44,43].

5 Conclusions

Computer simulations represent nowadays one of the key tools in the theoretical investigation of soft and biological matter. Despite their great success substantiated by increasing computer power, they are still limited in the number of particles and reachable timescales. A particular important class of systems involves water confined in small interlamellar regions, pores, cavities, and pockets of various proteins, which are in contact with a bulk water reservoir and hence at defined chemical potential. As one cannot afford simulating the entire bulk reservoir of water, special approaches are required in order to make progress for such systems.

In this chapter we discuss a recently introduced approach to deal with the problem of water between interacting surfaces in equilibrium with a bulk aqueous phase. We present the techniques for determining chemical potentials with high spatial resolution in detail. For the case of water, this is achieved by splitting its excess part into Lennard-Jones and electrostatic contributions. For best efficiency, we use a combination of test particle insertion and thermodynamic integration. Based on thermodynamic relations, we introduce the thermodynamic extrapolation technique, which enables us to evaluate the interaction pressure of the system at prescribed chemical potential. Using self-assembled monolayers as versatile surface models, we show basic features of the hydration repulsion, which universally acts between hydrophilic objects in water and discuss the significance of van der Waals interactions. While the hydration repulsion shows a monotonic decay with distance between soft surfaces, stiff surfaces exhibit oscillatory pressure profiles due to water layering. In the end we show that utilizing molecular modeling it is possible to reproduce the experimental results for the interaction between experimentally well-studied phospholipids bilayers. Such simulations enable for a deeper molecular insight into the mechanisms that

govern the interaction and into the phenomena present on the nanoscale. It hence enables us to test and refine available predictions from continuum theories and to circumvent limitations of simulation sizes and timescales.

Acknowledgments

We gratefully acknowledge funding from the DFG via SFB 765 and SFB 1112.

References

1. Roger H. French, V. Adrian Parsegian, Rudolf Podgornik, Rick F. Rajter, Anand Jagota, Jian Luo, Dilip Asthagiri, Manoj K. Chaudhury, Yet-ming Chiang, Steve Granick, Sergei Kalinin, Mehran Kardar, Roland Kjellander, David C. Langreth, Jennifer Lewis, Steve Lustig, David Wesolowski, John S. Wettlaufer, Wai-Yim Ching, Mike Finnis, Frank Houlihan, O. Anatole von Lilienfeld, Carel Jan van Oss, and Thomas Zemb, *Long range interactions in nanoscale science*, Reviews of Modern Physics, **82**, no. 2, 1887–1944, June 2010.
2. Daan Frenkel and Berend Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, Academic Press, Oct. 2001.
3. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, 1989.
4. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *Molecular dynamics with coupling to an external bath*, The Journal of Chemical Physics, **81**, no. 8, 3684–3690, Oct. 1984.
5. Shuichi Nosé, *A unified formulation of the constant temperature molecular dynamics methods*, The Journal of Chemical Physics, **81**, no. 1, 511–519, July 1984.
6. M. Parrinello and A. Rahman, *Polymorphic transitions in single crystals: A new molecular dynamics method*, Journal of Applied Physics, **52**, no. 12, 7182–7190, Dec. 1981.
7. H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *The missing term in effective pair potentials*, The Journal of Physical Chemistry, **91**, no. 24, 6269–6271, 1987.
8. Felix Sedlmeier, Dominik Horinek, and Roland R. Netz, *Spatial Correlations of Density and Structural Fluctuations in Liquid Water: A Comparative Simulation Study*, Journal of the American Chemical Society, **133**, no. 5, 1391–1398, 2011.
9. Bertrand Guillot, *A reappraisal of what we have learnt during three decades of computer simulations on water*, Journal of Molecular Liquids, **101**, no. 1–3, 219–260, Nov. 2002.
10. B. Widom, *Some Topics in the Theory of Fluids*, The Journal of Chemical Physics, **39**, no. 11, 2808–2812, Dec. 1963.
11. K.S. Shing and K.E. Gubbins, *The chemical potential in dense fluids and fluid mixtures via computer simulation*, Molecular Physics, **46**, no. 5, 1109–1128, Aug. 1982.
12. Siewert-Jan Marrink and Herman J. C. Berendsen, *Simulation of water transport through a lipid membrane*, The Journal of Physical Chemistry, **98**, no. 15, 4155–4168, 1994.
13. John G. Kirkwood, *Statistical Mechanics of Fluid Mixtures*, The Journal of Chemical Physics, **3**, no. 5, 300–313, May 1935.

14. Robert W. Zwanzig, *High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases*, The Journal of Chemical Physics, **22**, no. 8, 1420–1426, Aug. 1954.
15. Charles H Bennett, *Efficient estimation of free energy differences from Monte Carlo data*, Journal of Computational Physics, **22**, no. 2, 245–268, 1976.
16. Michael R. Shirts and John D. Chodera, *Statistically optimal analysis of samples from multiple equilibrium states*, The Journal of Chemical Physics, **129**, no. 12, 124105, Sept. 2008.
17. Chris Oostenbrink, Alessandra Villa, Alan E. Mark, and Wilfred F. Van Gunsteren, *A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6*, Journal of Computational Chemistry, **25**, no. 13, 1656–1676, 2004.
18. David Van Der Spoel, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E. Mark, and Herman J. C. Berendsen, *GROMACS: Fast, flexible, and free*, Journal of Computational Chemistry, **26**, no. 16, 1701–1718, 2005.
19. Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl, *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*, Journal of Chemical Theory and Computation, **4**, no. 3, 435–447, 2008.
20. Giovanni Bussi, Davide Donadio, and Michele Parrinello, *Canonical sampling through velocity rescaling*, The Journal of Chemical Physics, **126**, no. 1, 014101, Jan. 2007.
21. Andrij Trokhymchuk and José Alejandre, *Computer simulations of liquid/vapor interface in Lennard-Jones fluids: Some questions and answers*, The Journal of Chemical Physics, **111**, no. 18, 8510–8523, Nov. 1999.
22. Travis Oliphant, *A Bayesian perspective on estimating mean, variance, and standard-deviation from data*, All Faculty Publications, Dec. 2006.
23. Alexander Schlaich, "Modified GROMACS sources for TPI insertion into a slab; available on GitHub", 2014.
24. Naveen Michaud-Agrawal, Elizabeth J. Denning, Thomas B. Woolf, and Oliver Beckstein, *MDAnalysis: A toolkit for the analysis of molecular dynamics simulations*, Journal of Computational Chemistry, **32**, no. 10, 2319–2327, 2011.
25. Felix Sedlmeier, Dominik Horinek, and Roland R. Netz, *Nanoroughness, Intrinsic Density Profile, and Rigidity of the Air-Water Interface*, Physical Review Letters, **103**, no. 13, 136102, Sept. 2009.
26. Fernando A. Escobedo and Juan J. de Pablo, *Monte Carlo simulation of the chemical potential of polymers in an expanded ensemble*, The Journal of Chemical Physics, **103**, no. 7, 2703–2710, Aug. 1995.
27. Phil Attard, *Simulation of the chemical potential and the cavity free energy of dense hard-sphere fluids*, The Journal of Chemical Physics, **98**, no. 3, 2225–2231, Feb. 1993.
28. R. Delgado-Buscalioni, G. De Fabritiis, and P. V. Coveney, *Determination of the chemical potential using energy-biased sampling*, The Journal of Chemical Physics, **123**, no. 5, 054105, Aug. 2005.
29. Kevin B. Daly, Jay B. Benziger, Pablo G. Debenedetti, and Athanassios Z. Panagiotopoulos, *Massively parallel chemical potential calculation on graphics processing units*, Computer Physics Communications, **183**, no. 10, 2054–2062, 2012.

30. Emanuel Schneck and Roland R. Netz, *From simple surface models to lipid membranes: Universal aspects of the hydration interaction from solvent-explicit simulations*, Current Opinion in Colloid & Interface Science, **16**, no. 6, 607–611, Dec. 2011.
31. D. Bratko, R. A. Curtis, H. W. Blanch, and J. M. Prausnitz, *Interaction between hydrophobic surfaces with metastable intervening liquid*, The Journal of Chemical Physics, **115**, no. 8, 3873–3877, Aug. 2001.
32. Tomohiro Hayashi, Alexander J Pertsin, and Michael Grunze, *Grand canonical Monte Carlo simulation of hydration forces between nonorienting and orienting structureless walls*, The Journal of Chemical Physics, **117**, no. 13, 6271–6280, Oct. 2002.
33. Emanuel Schneck, Felix Sedlmeier, and Roland R. Netz, *Hydration repulsion between biomembranes results from an interplay of dehydration and depolarization*, Proceedings of the National Academy of Sciences, **109**, no. 36, 14405–14409, Sept. 2012.
34. Kazi A. Motakabbir and M. Berkowitz, *Isothermal compressibility of SPC/E water*, The Journal of Physical Chemistry, **94**, no. 21, 8359–8362, 1990.
35. Helena L. Pi, Juan L. Aragones, Carlos Vega, Eva G. Noya, Jose L.F. Abascal, Miguel A. Gonzalez, and Carl McBride, *Anomalies in water as obtained from computer simulations of the TIP4P/2005 model: density maxima, and density, isothermal compressibility and heat capacity minima*, Molecular Physics, **107**, no. 4-6, 365–374, Feb. 2009.
36. Rana A. Fine and Frank J. Millero, *Compressibility of water as a function of temperature and pressure*, The Journal of Chemical Physics, **59**, no. 10, 5529–5536, Nov. 1973.
37. Matej Kanduč, Emanuel Schneck, and Roland R. Netz, *Attraction between hydrated hydrophilic surfaces*, Chemical Physics Letters, **610–611**, 375–380, Aug. 2014.
38. Tom Darden, Darrin York, and Lee Pedersen, *Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems*, The Journal of Chemical Physics, **98**, no. 12, 10089–10092, June 1993.
39. Ulrich Essmann, Lalith Perera, Max L. Berkowitz, Tom Darden, Hsing Lee, and Lee G. Pedersen, *A smooth particle mesh Ewald method*, The Journal of Chemical Physics, **103**, no. 19, 8577–8593, Nov. 1995.
40. Alexander Schlaich, Bartosz Kowalik, Matej Kanduč, Emanuel Schneck, and Roland R. Netz, *Physical mechanisms of the interaction between lipid membranes in the aqueous environment*, Physica A: Statistical Mechanics and its Applications, **418**, 105–125, Jan. 2015.
41. V. Adrian Parsegian, *Van der Waals Forces: A Handbook for Biologists, Chemists, Engineers, and Physicists*, Cambridge University Press, Nov. 2005.
42. H. C. Hamaker, *The London—van der Waals attraction between spherical particles*, Physica, **4**, no. 10, 1058–1072, 1937.
43. VA Parsegian, N. Fuller, and RP Rand, *Measured work of deformation and repulsion of lecithin bilayers*, Proceedings of the National Academy of Sciences, **76**, no. 6, 2750–2754, 1979.
44. LJ Lis, M. McAlister, N. Fuller, RP Rand, and VA Parsegian, *Interactions between neutral phospholipid bilayer membranes.*, Biophysical journal, **37**, no. 3, 657, 1982.
45. J.N. Israelachvili and G.E. Adams, *Measurement of forces between two mica surfaces in aqueous electrolyte solutions in the range 0–100 nm*, Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases, **74**, 975–1001, 1978.

46. RP Rand and VA Parsegian, *Hydration forces between phospholipid bilayers*, Biochimica et biophysica acta, **988**, no. 3, 351–376, 1989.
47. R E Pagano and J N Weinstein, *Interactions of Liposomes with Mammalian Cells*, Annual Review of Biophysics and Bioengineering, **7**, no. 1, 435–468, 1978.
48. S. Marčelja and N. Radić, *Repulsion of interfaces due to boundary water*, Chemical Physics Letters, **42**, no. 1, 129–130, Aug. 1976.
49. Matej Kanduč, Alexander Schlaich, Emanuel Schneck, and Roland R. Netz, *Hydration repulsion between membranes and polar surfaces: Simulation approaches versus continuum theories*, Advances in Colloid and Interface Science, **208**, 142–152, June 2014.
50. Jacob Israelachvili and Hakan Wennerström, *Role of hydration and water structure in biological and colloidal interactions*, Nature, **379**, no. 6562, 219–225, Jan. 1996.
51. Matej Kanduč, Emanuel Schneck, and Roland R. Netz, *Hydration Interaction between Phospholipid Membranes: Insight into Different Measurement Ensembles from Atomistic Molecular Dynamics Simulations*, Langmuir, **29**, no. 29, 9126–9137, July 2013.
52. Gregor Cevc and Derek Marsh, *Phospholipid bilayers: physical principles and models*, Wiley, May 1987.
53. C. Tanford, *The hydrophobic effect and the organization of living matter*, Science, **200**, no. 4345, 1012–1018, June 1978.
54. Reinhard Lipowsky, *The conformation of membranes*, Nature, **349**, no. 6309, 475–481, Feb. 1991.
55. O. Berger, O. Edholm, and F. Jähnig, *Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature*, Biophysical Journal, **72**, no. 5, 2002–2013, May 1997.

Classical Density Functional Theory to Tackle Solvation in Molecular Liquids

Guillaume Jeanmairet, Maximilien Levesque, Volodymyr Sergiievskyi, and Daniel Borgis

École Normale Supérieure, Département de Chimie
UMR 8640 CNRS-ENS-UPMC, 24, rue Lhomond, 75005 Paris, France

E-mail: {*daniel.borgis,maximilien.levesque*}*@ens.fr*

We present a brief review of the classical density functional theory of atomic and molecular fluids. We focus on the application of the theory to the determination of the solvation properties of arbitrary molecular solutes in arbitrary molecular solvent. This includes the prediction of the solvation free energies, as well as the characterization of the microscopic, three-dimensional solvent structure.

1 Introduction

The determination of the solvation free-energy of molecular solutes in molecular solvents is a problem of primary importance in physical chemistry and biology. From a theoretical point of view, two extreme strategies can be found in the literature. A standard route consists in using molecular simulation techniques such as molecular dynamics (MD) or Monte-Carlo (MC), with an explicit molecular solvent. There are a number of well-established statistical mechanics techniques to estimate absolute or relative free-energies by molecular simulations, for example thermodynamic integration methods based on umbrella sampling[1,2], or generalized constraints[3]. In any case, the precise estimation of free-energies by computer simulation remains extremely costly; it requires to consider a sufficiently large number of solvent molecules around the molecular solute and, for this large system, to average a "generalized force" over many microscopic solvent configurations, and this for a lot of different points along the reversible thermodynamic integration path.

Another class of methods, known as implicit solvent models[4], relies on the assumption that the macroscopic laws remain valid at a microscopic level, and that solvation free energies can be computed by combining a dielectric continuum description of the solvent outside the solute core and a simple solvent-accessible surface area expression for the non-electrostatic contributions[5]. For the electrostatic part, the stationary Poisson-Boltzmann equation can be solved for the electrostatic potential using sharp definitions of the dielectric boundaries and various efficient numerical techniques, making it possible to handle very large biomolecular systems[6]. Density functional methods based on the minimization of polarization density[7] have been introduced too. There are serious limitations however to a continuum dielectric approach, and first of all the validity of the macroscopic electrostatic laws at microscopic distances, the neglect of the molecular nature of the solvent, and the ambiguous definition of all the non-electrostatic energetic contributions, such as hydrophobicity. Macroscopic approaches to hydrophobicity that mixes consistently with the Poisson-Boltzmann description are presently developed.[8,9]

Beyond continuum descriptions, it is desirable however to devise and employ implicit solvent methods which (i) are able to cope with the molecular nature of the solvent, but without considering explicitly all its instantaneous microscopic degrees of freedom, and (ii) can provide solvation properties at a modest computer cost compared to explicit simulations. Such methods should rely on the theory of molecular liquids that has been developed in the second half of the last century and lie by now in classical textbooks[10–12]. Among possible approaches one should mention molecular integral equation theories in the reference interaction site (RISM)[13–16] or molecular[17–23], or mixed[24,25] picture, Gaussian field theories[26,27], the density functional theory (DFT) of molecular liquids[10,28–39], or, finally, field theoretical approaches to dipolar solvent-ions mixtures, that lead to a generalization of the Poisson-Boltzmann equation accounting for particle size and dielectric saturation[40–44]. Note that, close to our purpose, a 3D-version of the RISM equations has been developed recently to describe the solvation of objects of complex shape[45–50].

Our focus here is classical density functional theory (DFT), and eventually a molecular version of it that we will call Molecular Density Functional Theory (MDFT). The basic theoretical principles of classical DFT can be found in the seminal paper by B. Evans[28] and subsequent excellent reviews by him[28–30] and other authors[51]. The advent in the late 1980's of a quasi-exact DFT for inhomogeneous hard sphere mixtures, the fundamental measure theory[52–57], has promoted recently a great deal of applications to atomic-like fluids in bulk or confined conditions or at interfaces. Classical "atomic" DFT can be considered nowadays as a method of choice for many chemical engineering problems[58,59]. Much less applications exist for molecular fluids, for which solvent orientations should be considered. The description has been generally limited to generic dipolar solvents[60,61] or dipolar solvent/ions mixtures[33–36]; such approach may be considered already as "civilized" compared to primitive continuum models[35]. We have proposed recently an extension of MDFT to arbitrary fluid/solvents in the precise goal of describing the solvation of three-dimensional molecular object in arbitrary solvents.[37,39,38,62–70] A RISM-based DFT approach of molecular solvation has been developed recently too[71].

The outline of the present review is as follows. We first recall the basic principles of cDFT for atomic-like fluids. We then describe the particular but fundamental case of the hard-sphere fluid, and the associated fundamental measure theory (FMT), focusing on a "scalar" formulation due to Kierlik and Rosinberg[53,54], instead of the standard "vectorial" version introduced initially by Rosenfeld[52]. We then turn to Lennard-Jones fluids, for which the HS fluid can be used as a reference in various ways to construct a functional. The last section will be devoted to molecular solvent, modeled by rigid polyatomic molecules with an orientation. The discussion will focus on a model dipolar solvent, the Stockmayer fluid, and then extend to realistic models of polar liquids such as acetonitrile and water.

2 The Case of Atomic Fluids

2.1 General formulation

In this section we begin with recalling the basis of the density functional theory of liquids, and discussing the general problem of a molecular solvent submitted to an external field. In the applications we have in mind, the external field will be created by a molecular solute of arbitrary shape dissolved at infinite dilution in the solvent. The individual solvent

molecules will be considered later as rigid bodies described by their position $\mathbf{r}$ and orientation ω. In this section we restrict the discussion to atomic or pseudo-atomic solvents (such as CCl_4) modeled by spherical particles for which only the position $\mathbf{r}$ matters.

The grand potential density functional for a fluid having an inhomogeneous density $\rho(\mathbf{r})$ in the presence of an external field $V_{ext}(\mathbf{r})$ can be defined as[28,29],

$$\Omega[\rho] = F[\rho] - \mu_s \int \rho(\mathbf{r}) d\mathbf{r}, \tag{1}$$

where $F[\rho]$ is the Helmholtz free energy functional and μ_s is the chemical potential. The grand potential can be evaluated relatively to a reference homogeneous fluid having the same chemical potential μ_s and particle density ρ_0

$$\Omega[\rho] = \Omega[\rho_0] + \mathcal{F}[\rho]. \tag{2}$$

Following the general theoretical scheme introduced by Evans[28,29,10,72], the density functional $\mathcal{F}[\rho]$ can be split into three contributions: an ideal term, an external potential term and an excess free-energy term accounting for the intrinsic interactions within the fluid,

$$\mathcal{F}[\rho] = \mathcal{F}_{id}[\rho] + \mathcal{F}_{ext}[\rho] + \mathcal{F}_{exc}[\rho], \tag{3}$$

with the following expressions of the first two terms

$$\mathcal{F}_{id}[\rho]{=}k_B T \int d\mathbf{r}_1 \left[\rho(\mathbf{r}_1) \ln\left(\frac{\rho(\mathbf{r}_1)}{\rho_0} \right) - \rho(\mathbf{r}_1) + \rho_0 \right], \tag{4}$$

$$\mathcal{F}_{ext}[\rho] = \int d\mathbf{r}_1 \, V_{ext}(\mathbf{r}_1) \rho(\mathbf{r}_1). \tag{5}$$

There are several ways to arrive to an exact expression of the excess free-energy, i.e. using an adiabatic perturbation of the pair potential (the so-called adiabatic connection route in electronic DFT), of the external potential, or of the density itself. If the latest route is chosen, one can define $\mathcal{F}_{exc}[\rho]$ as:[28]

$$\mathcal{F}_{exc}[\rho]{=}k_B T \iint d\mathbf{r}_1 d\mathbf{r}_2 \, C(\mathbf{r}_1, \mathbf{r}_2) \, \Delta\rho(\mathbf{r}_1) \Delta\rho(\mathbf{r}_2), \tag{6}$$

with $\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_0$. The function $C(\mathbf{r}_1, \mathbf{r}_2)$ is still a functional of $\rho(\mathbf{r})$ defined by

$$C(\mathbf{r}_1, \mathbf{r}_2) = \int_0^1 d\alpha \, (\alpha - 1) \; c^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho_\alpha]), \tag{7}$$

where $c^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho_\alpha])$ is the two particle direct correlation function, i.e. by definition the second order derivative of the excess free-energy with respect to density, evaluated at the intermediate density $\rho_\alpha(\mathbf{r}) = \rho_0 + \alpha \Delta\rho(\mathbf{r})$. Eqs 6-7 follows naturally when expressing the functional from the knowledge of its second-derivatives[28].

The equilibrium condition reads

$$\left. \frac{\delta\Omega[\rho]}{\delta\rho} \right|_{\rho=\rho_{eq}} = 0 \qquad \Longrightarrow \left. \frac{\delta\mathcal{F}[\rho]}{\delta\rho} \right|_{\rho=\rho_{eq}} = 0. \tag{8}$$

When minimizing the density functional $\mathcal{F}[\rho]$ with respect to $\rho(\mathbf{r})$, the value at the minimum is the difference of the solvent grand potential with and without the solute, and thus

the solute solvation free-energy. The associated density $\rho_{eq}(\mathbf{r})$ is the equilibrium inhomogeneous density.

The functional defined by eqs 3-6 is formally exact but the inhomogeneous direct correlation functions entering the definition of the excess term are indeed unknown. However, simple approximations can be proposed for this quantity. The most natural one consists in expanding the inhomogeneous direct correlation function $c^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho_\alpha])$ around $\alpha = 0$, that is, around the homogeneous density ρ_0:

$$c^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho_\alpha]) = c^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho_0]) + \alpha \int d\mathbf{r}_3 \, c^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; [\rho_0]) \, \Delta\rho(\mathbf{r}_3) + ... \quad (9)$$

Such expression involves the two, three, n-particle direct correlation functions of the homogeneous fluid[10]. The first term is the (two-body) direct correlation function (DCF) of the homogeneous solvent, that depends on $r_{12} = |\mathbf{r}_2 - \mathbf{r}_1|$, and can be thus denoted as $c_S(r_{12}; \rho_0)$ (S for spherical component, preparing ourselves to non-spherical solvents).

Using eq. 7, the excess term can thus be written as

$$\mathcal{F}_{exc}[\rho] = -\frac{k_B T}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 \, c_S(r_{12}; \rho_0) \, \Delta\rho(\mathbf{r}_1) \Delta\rho(\mathbf{r}_2) + \mathcal{F}_B[\rho], \quad (10)$$

where we have defined the bridge functional

$$\mathcal{F}_B[\rho] = -\frac{k_B T}{6} \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \, c^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3; [\rho_0]) \, \Delta\rho(\mathbf{r}_1) \, \Delta\rho(\mathbf{r}_2) \, \Delta\rho(\mathbf{r}_3) + O(\Delta\rho^4), \quad (11)$$

which thus starts with a cubic term in $\Delta\rho$. Setting $\mathcal{F}_B[\rho] = 0$ correspond to the so-called homogeneous reference fluid (HRF) approximation. It can be shown to be equivalent to the hypernetted chain (HNC) approximation in integral equation theories[30]. The input of the theory is the direct correlation function of the pure solvent, which can be extracted from simulation or experimental data by measuring the total correlation function $h_S(r) = g(r) - 1$ and solving subsequently the Ornstein-Zernike equation, i.e. in Fourier space:

$$1 - \rho_0 c_S(k) = (1 + \rho_0 h_S(k))^{-1} = \chi_n^{-1}(k). \quad (12)$$

$\chi_n(r)$ is the structure factor, or the density susceptibility, measuring density-density correlations at a given distance in the fluid. The excess free energy can thus be expressed also in terms of the susceptibility

$$\mathcal{F}_{exc}[\rho] = \frac{k_B T}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 \, \chi_n^{-1}(r_{12}) \, \Delta\rho(\mathbf{r}_1) \Delta\rho(\mathbf{r}_2) - \frac{k_B T}{2\rho_0} \int d\mathbf{r} \, \Delta\rho(\mathbf{r})^2 + \mathcal{F}_B[\rho]. \quad (13)$$

2.2 Fundamental Measure Theory for the hard-sphere fluid: Scalar versus Vectorial Formulation

We focus here on the particular case of the hard-sphere fluid and describe briefly the fundamental measure theory (FMT) introduced by Rosenfeld[52] and Kierlik and Rosinberg[53]. Although the theory is valid for arbitrary hard-sphere mixtures, we consider a one-component HS fluid composed of hard spheres of radius R and at a bulk density ρ_0. The fluid is subjected to an external perturbation, for example a solid interface or a molecular solute of arbitrary shape embedded in the fluid, that creates a position-dependent external potential

$V_{ext}(\mathbf{r})$ and thus an inhomogeneous density $\rho(\mathbf{r})$. The excess functional of eq. 6 can be written as

$$\mathcal{F}_{exc}[\rho] = F_{exc}^{HS}[\rho] - F_{exc}^{HS}[\rho_0] - \mu_{exc}^{HS} \int d\mathbf{r}\,(\rho(\mathbf{r}) - \rho_0)\,, \tag{14}$$

where $F_{exc}^{HS}(\rho)$ is the excess Helmholtz free-energy functional for the hard-sphere fluid and μ_{exc}^{HS} is the bulk excess chemical potential defined by

$$\mu_{exc}^{HS} = \frac{\delta \mathcal{F}_{exc}^{HS}[\rho]}{\delta \rho}|_{\rho=\rho_0}, \tag{15}$$

so that obviously $\frac{\delta \mathcal{F}_{exc}[\rho]}{\delta \rho}|_{\rho=\rho_0} = 0$. In the FMT introduced by Rosenfeld[52], the excess functional for the hard-sphere fluid can be written in terms of a set of N_w weighted densities, $n_\alpha(\mathbf{r})$:

$$\mathcal{F}_{exc}[\rho(\mathbf{r})] = k_B T \int d\mathbf{r}\,\Phi(\{n_\alpha(\mathbf{r})\}) \tag{16}$$

with

$$n_\alpha(\mathbf{r}) = \int d\mathbf{r}'\,\rho(\mathbf{r}')\,\omega_\alpha(\mathbf{r}-\mathbf{r}') = \rho(\mathbf{r}) \star \omega_\alpha(\mathbf{r}), \tag{17}$$

where $\omega_\alpha(\mathbf{r})$ are geometrical weight functions to be defined below and $\star$ indicates the convolution of the microscopic densities by those weight functions.

In the original Rosenfeld's derivation there are four scalar weight functions, $\omega_\alpha^i(\mathbf{r}), \alpha = 0,1,2,3$, and two vectorial ones $\vec{\omega}_1(\mathbf{r}), \vec{\omega}_2(\mathbf{r})$ that are defined by

$$\omega_3(\mathbf{r}) = \Theta(R-r) \tag{18}$$

$$\omega_2(\mathbf{r}) = 4\pi R\,\omega_1(\mathbf{r}) = 4\pi R^2\,\omega_0(\mathbf{r}) = \delta(R-r) \tag{19}$$

$$\vec{\omega}_2(\mathbf{r}) = 4\pi R\,\vec{\omega}_1(\mathbf{r}) = \frac{\mathbf{r}}{r}\,\delta(R-r). \tag{20}$$

$\Theta(r)$ denotes the Heaviside function and $\delta(r)$ the Dirac distribution. The excess free-energy density Φ derived by Rosenfeld for Eq. 16 is a function of the four position-dependent weighted densities, $n_\alpha(\mathbf{r}), \alpha = 0,1,2,3$, and of the two vectorial ones, $\vec{n}_1(\mathbf{r}), \vec{n}_2(\mathbf{r})$, which generates in the homogeneous limit the Percus-Yevick equation of state for hard-sphere mixtures. Starting from the generalization of the Carnahan-Starling (CS) equation of state to mixtures (namely the Mansoori-Carnahan-Starling-Leland equation (MCSL)) instead of PY, Roth et al[55] and Wu et al[56] were later able to obtain a modified expression based on the same definition of the weighted densities (either called white-bear (WB) version or modified FMT version (MFMT)). This modified version of FMT takes advantage of the fact that the CS expression provides one a better equation of state that PY.

Ten years before those latest developments, Kierlick and Rosinberg were able to derive an alternative version of FMT which involves only four scalar weight functions $\omega_\alpha(\mathbf{r}), \alpha = 0,1,2,3$.[53,54]. The last two weights are identical to Eq. 18-19, whereas the first two ones are given by

$$\omega_1(\mathbf{r}) = \frac{1}{8\pi}\delta'(R-r) \tag{21}$$

$$\omega_0(\mathbf{r}) = \frac{1}{8\pi}\,\delta''(R-r) + \frac{1}{2\pi r}\,\delta'(R-r) \tag{22}$$

Those weight functions appear naturally in the derivation as the inverse Fourier transforms of

$$\begin{aligned}
\omega_3(k) &= \frac{4\pi}{k^3}(\sin(kR) - kR\cos(kR)) \\
\omega_2(k) &= \frac{4\pi R}{k}\sin(kR) \\
\omega_1(k) &= \frac{1}{2k}(\sin(kR) + kR\cos(k)) \\
\omega_0(k) &= \cos(kR) + \frac{kR}{2}\sin(kR).
\end{aligned} \tag{23}$$

Although the main part of the papers by Kierlik and Rosinberg relies on a PY expression for the excess free energy density

$$\Phi^{\mathrm{PY}}[\{n_\alpha\}] = -n_0\ln(1-n_3) + \frac{n_1 n_2}{1-n_3} + \frac{1}{24\pi}\frac{n_2^3}{(1-n_3)^2}, \tag{24}$$

the authors do mention in their conclusion that a CS (more precisely MCSL) expression could be used instead

$$\Phi^{\mathrm{CS}}[\{n_\alpha\}] = \left(\frac{1}{36\pi}\frac{n_2^3}{n_3^2} - n_0\right)\ln(1-n_3) + \frac{n_1 n_2}{1-n_3} + \frac{1}{36\pi}\frac{n_2^3}{(1-n_3)^2 n_3}. \tag{25}$$

They point out the fact that this expression is more precise than the PY one, but using it while keeping the expression of the weights unchanged leads to thermodynamic inconsistencies; those inconsistencies are indeed present in the WB or MFMT formulations too. There is clearly a trade off to be made between precision and theoretical consistency. It was later shown by Phan et al. that the Kierlik and Rosinberg's approach is mathematically equivalent to the original vectorial version.[73] On a practical point of view, however, and especially in the perspective of 3D applications, the KR formulation is advantageous with respect to the Rosenfeld's formulation since the number of *independent* weighted densities is reduced from 5 to 4 for the one component system, and from $(10 + N_s)$ to $(4 + N_s)$ for a mixture of N_s components, with $N_s \geq 2$; thus from 12 to 6 for a binary mixture. An efficient numerical implementation in three-dimensions of the Kierlik-Rosinberg FMT functional is detailed in Ref.[65]. The numerical efficiency of the algorithms, in terms of convergence rate and system size dependency, is briefly illustrated in Fig. 1.

2.3 The Lennard-Jones fluid

Building the thermodynamics of the Lennard-Jones fluid by taking the suitable hard-sphere fluid as a reference and building in the attractive interaction as a perturbation is indeed a classics in liquid state theory and is at the basis of the Van der Waals theory of fluids. When coming to functionals, this idea can be declined in several variants, the most natural one being to use a FMT functional for the repulsive part and a mean-field approximation (or mean spherical approximation, MSA) for the attractive part[30]. Along the lines given above, another route is to approximate the bridge functional in eq. 10 by a hard sphere

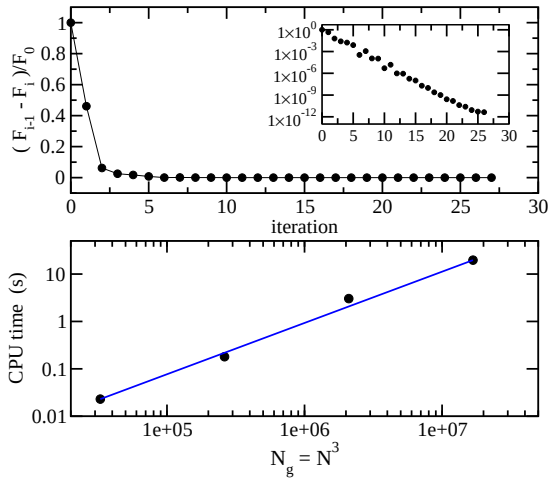


Figure 1: Top: Typical plot of the free energy difference between two successive steps (normalized by the initial energy) versus minimization-step number (Here a benzene molecule in a one-component HS reference fluid modeling SPC water). The inlet represents the same with a logarithmic scale in ordinates. Bottom: CPU time per minimization step versus number of 3D-grid points. The circle correspond in increasing order to N= 32, 64, 128, and 256.

bridge functional, introduced by Rosenfeld as a universal bridge function[74,75]

$$\begin{aligned}\mathcal{F}[\rho(\mathbf{r})] =& k_B T \int d\mathbf{r}\, \left[\rho(\mathbf{r}) \ln\left(\frac{\rho(\mathbf{r})}{\rho_0}\right) - \rho(\mathbf{r}) + \rho_0\right] + \int d\mathbf{r}\, V_{ext}(\mathbf{r})\rho(\mathbf{r}) \\ & - \frac{k_B T}{2} \int d\mathbf{r}_1 d\mathbf{r}_2\, c_S(r_{12};\rho_0)\Delta\rho(\mathbf{r}_1)\,\Delta\rho(\mathbf{r}_2) + \mathcal{F}_B^{HS}[\rho],\end{aligned} \tag{26}$$

where

$$\begin{aligned}\mathcal{F}_B^{HS}[\rho(\mathbf{r})] =& F_{exc}^{HS}[\rho(\mathbf{r})] - F_{exc}^{HS}[\rho_0] - \mu_{exc}^{HS} \int d\mathbf{r}\Delta\rho(\mathbf{r}) \\ & + \frac{k_B T}{2} \int d\mathbf{r}_1 d\mathbf{r}_2\, c_S^{HS}(r_{12};\rho_0)\Delta\rho(\mathbf{r}_1)\,\Delta\rho(\mathbf{r}_2).\end{aligned} \tag{27}$$

The first three terms represent the one-component hard-sphere KR-FMT excess functional defined in the previous section and the associated chemical potential yielding equilibrium at $\rho(\mathbf{r}) = \rho_0$. The fourth term involves the direct correlation function of the HS fluid at the same density, i.e

$$c_S^{HS}(|\mathbf{r}_1 - \mathbf{r}_2|;\rho_0) = -\frac{\delta^2 \mathcal{F}_{exc}^{HS}[\rho]}{\delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_2)}|_{\rho(\mathbf{r})=\rho_0}. \tag{28}$$

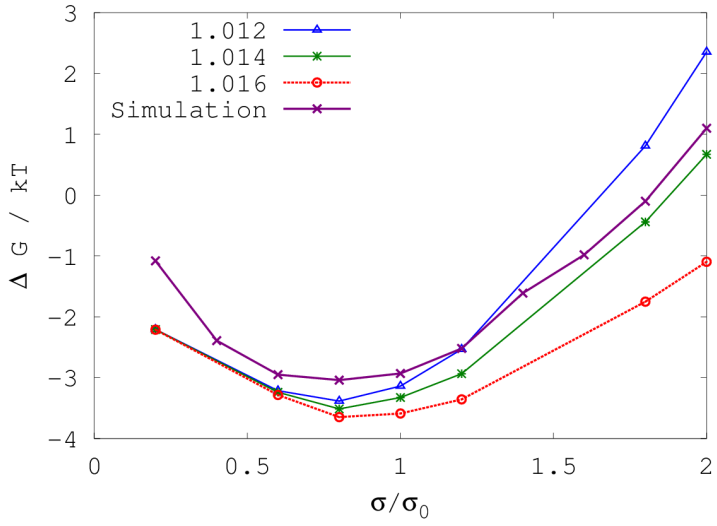


Figure 2: Solvation free-energy obtained by DFT using the HS bridge functional of 27 with different HS diameters, compared to the Monte-Carlo results of Lazaridis[77].

This function can be easily obtained in Fourier space as

$$c_S^{HS}(k;\rho_0) = -\sum_{\alpha,\beta} \frac{\partial^2 \Phi(\{n_\gamma^0\})}{\partial n_\alpha \partial n_\beta}\, \omega_\alpha(k)\omega_\beta(k), \tag{29}$$

where $\{n_\gamma^0\}$ represent the weighted densities for a uniform fluid of density ρ_0 and the $\omega_{\alpha,\beta}(k)$ are the weights of eq. 23. The second derivatives have to be taken for the PY or CS functions of eqs 24 or 25. Note that defined as in eq. 27, $\mathcal{F}_B^{HB}[\rho(\mathbf{r})]$ carries an expansion in $\Delta\rho$ of order 3 and higher that corrects the second order expansion of the excess free energy in eq. 26. The excess functional can also be re-expressed as

$$\begin{aligned}\mathcal{F}_{exc}[\rho] =& F_{exc}^{HS}[\rho] - F_{exc}^{HS}[\rho_0] - \mu_{exc}^{HS} \int d\mathbf{r} \Delta\rho(\mathbf{r}) \\ &- \frac{k_B T}{2} \int d\mathbf{r}_1 d\mathbf{r}_2\, c_S^{att}(r_{12};\rho_0)\, \Delta\rho(\mathbf{r}_1)\Delta\rho(\mathbf{r}_2),\end{aligned} \tag{30}$$

where we have defined the "attractive" DCF by

$$c_S^{att}(r_{12};\rho_0) = c_S(r_{12};\rho_0) - c_S^{HS}(r_{12};\rho_0). \tag{31}$$

Eq. 30 is the basis of the first order mean-spherical approximation (FMSA) theory developed by Tang[76].

We show here how this FMSA theory works for our purpose: the prediction of solvation properties of dissolved molecular objects. In Fig. 2, we compare the solvation free energy of a LJ sphere of increasing diameter in a LJ fluid with $\rho^* = 0.85, T^* = 0.88$, as computed by Monte-Carlo simulations by Lazaridis[77], to the results obtained by DFT minimization

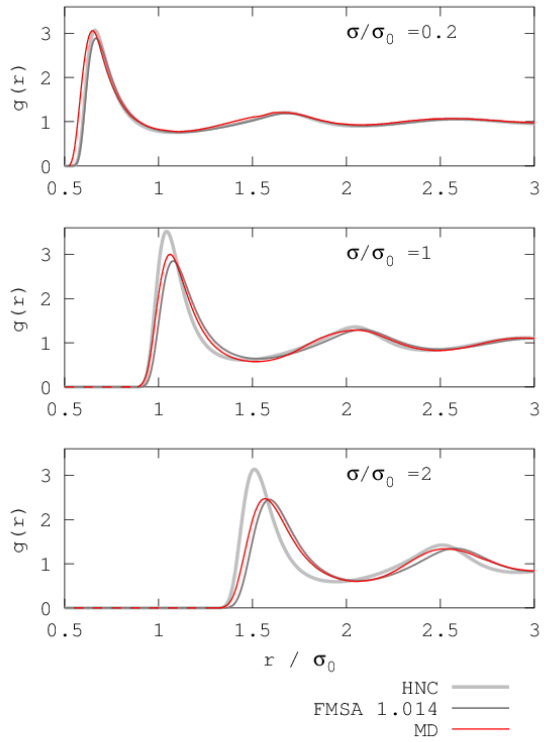


Figure 3: Reduced solvent density around LJ solutes of different diameters, using the HNC approximation, or adding a hard-sphere bridge functional with $d = 1.014\sigma$.

with different HS diameters, d. It can be seen that the results are extremely sensitive to the choice of d, and that the best agreement is obtained for $d = 1.014\sigma$ (indeed close to 1, that would be the initial guess value). For that value, we have plotted in Fig. 3 the microscopic solvent density, $g(r) = \rho(r)/\rho_0$, obtained for solute of different sizes by direct simulation, or by DFT in the HNC or FMSA approximation. It can be seen that the addition of hard-sphere bridge in FMSA greatly improve the results compared to the HRF (or HNC) approximation and yields a correct structure.

3 The Case of Molecular Fluids: Molecular Density Functional Theory (MDFT)

3.1 General formulation

The solvent molecules now carry a molecular structure that is described by a collection of distributed atomic interaction sites. The theory is formulated in the molecular picture in which each solvent molecule is considered as a rigid body and characterized by its position, $\mathbf{r}$ (e.g. the position of center of mass), and by its orientation, ω, defined by the three Euler angles $\omega = (\theta, \phi, \psi)$. Thus, in the presence of an external perturbation, the solvent is

now characterized by an inhomogeneous position and orientation density $\rho(\mathbf{r},\omega)$. The solute, as the solvent, is described in microscopic details by a molecular non-polarizable "force-field" involving atomic Lennard-Jones and partial charges parameters. Given that the solute is fixed and defined by the position, $\mathbf{R}_i$, of its different atomic sites, the external potential is defined by

$$V_{ext}(\mathbf{r},\omega) = \sum_{i\in solute}\sum_{j\in solvent} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}, \qquad (32)$$

where ϵ_{ij} and σ_{ij} are the Lennard-Jones parameters between solute site i and solvent site j, and q_i and q_j are the partial charges carried by those sites. The relative site-site vectors are function of the solvent molecule position and orientation and defined as $\mathbf{r}_{ij} = \mathbf{r} + \mathbf{R}(\omega)\mathbf{s}_j - \mathbf{R}_i$, where $\mathbf{s}_j$ denotes the site positions in the molecular frame and $\mathbf{R}(\omega)$ is the rotation matrix associated to ω.

The same density functional as in eqs 1-3 can be written for $\rho(\mathbf{r},\omega)$, with an ideal, external, and excess part:

$$\begin{aligned}
\mathcal{F}_{id}[\rho] =& k_B T \int d\mathbf{r}d\omega \left[\rho(\mathbf{r},\omega)\ln\left(\frac{8\pi^2\rho(\mathbf{r},\omega)}{n_0}\right) - \rho(\mathbf{r},\omega) + \frac{n_0}{8\pi^2}\right] \\
\mathcal{F}_{ext}[\rho] =& \int d\mathbf{r}d\omega\, V_{ext}(\mathbf{r},\omega)\rho(\mathbf{r},\omega) \\
\mathcal{F}_{exc}[\rho] =& -\frac{1}{2}k_B T \iint d\mathbf{r}_1 d\mathbf{r}_2 d\omega_1 d\omega_2 \Delta\rho(\mathbf{r}_1,\omega_1)\, c(\mathbf{r}_1-\mathbf{r}_2,\omega_1,\omega_2')\,\Delta\rho(\mathbf{r}_2,\omega_2) \\
&+ \mathcal{F}_B[\rho(\mathbf{r},\omega)],
\end{aligned} \qquad (33)$$

where $\Delta\rho(\mathbf{r},\omega) = \rho(\mathbf{r},\omega) - n_0/8\pi^2$, n_0 being the particle number density of the reference bulk fluid. The first term represents the Homogeneous Reference Fluid approximation (or HNC approximation) where the excess free-energy density is written in terms of the angular-dependent direct correlation of the *pure* solvent. The second term represents the unknown correction to that term (or Bridge functional) that, again, can be expressed as of a systematic expansion of the solvent correlations in terms of the three-body, .. n-body terms direct correlation functions.

3.2 The Stockmayer solvent

To test and illustrate the theory, we start from the simplest conceivable model of dipolar solvent, the Stockmayer model, characterized by a single Lennard-Jones center with parameters σ_s, ϵ_s and a dipole $\boldsymbol{\mu}_s = \mu\boldsymbol{\omega}$, where $\boldsymbol{\omega}$ is the unitary orientational vector of the molecule –which here replaces the orientation noted ω above. The parameters are selected to make the model look like water (similar density, $n_0 = 0.033$ particles/Å^3, particle size, $\sigma_s = 3$Å, and molecular dipole, $\mu = 1.85\,D$) although not tasting quite as water (no hydrogen bond in the model!). For such model, the external potential can be written as

$$V_{ext}(\mathbf{r},\boldsymbol{\omega}) = \Phi_{LJ}(\mathbf{r}) - \mu\mathbf{E}_q(\mathbf{r})\cdot\boldsymbol{\omega} \qquad (34)$$

with

$$\Phi_{LJ}(\mathbf{r}) = \sum_{i=1}^{M} 4\epsilon_{si}\left[\left(\frac{\sigma_{si}}{|\mathbf{r}-\mathbf{R}_i|}\right)^{12} - \left(\frac{\sigma_{si}}{|\mathbf{r}-\mathbf{R}_i|}\right)^{6}\right] \tag{35}$$

$$\mathbf{E}_q(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\sum_{i=1}^{M}\frac{q_i(\mathbf{r}-\mathbf{R}_i)}{|\mathbf{r}-\mathbf{R}_i|^3}. \tag{36}$$

It is also argued in Refs[17,19] that the c-functions can be expanded onto a rotational invariants basis set keeping, to a good approximation approximation, the same order as the intermolecular potential, namely

$$c(\mathbf{r}_{12},\boldsymbol{\omega}_1,\boldsymbol{\omega}_2) = c_S(r_{12}) + c_\Delta(r_{12})\,\Phi^{110}(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2) + c_D(r_{12})\,\Phi^{112}(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2), \tag{37}$$

where

$$\begin{aligned}\Phi^{110} &= \boldsymbol{\omega}_1\cdot\boldsymbol{\omega}_2,\\ \Phi^{112} &= 3\,(\boldsymbol{\omega}_1\cdot\hat{\mathbf{r}}_{12})\,(\boldsymbol{\omega}_2\cdot\hat{\mathbf{r}}_{12}) - \boldsymbol{\omega}_1\cdot\boldsymbol{\omega}_2\end{aligned} \tag{38}$$

represents the two first non-isotropic spherical invariants. The three components c_S, c_Δ, c_D of $c(\mathbf{r}_{12},\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)$ can be obtained from the the corresponding components of the total correlation function, $h(\mathbf{r}_{12},\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)$, by inversion of the angular-dependent OZ equation. The total correlation function itself can be measured by using, e.g., MD simulations. We have used here the original Wertheim's notations with subscripts S, Δ, and D for the different h- and c-components. In this approximation, it was shown in Refs[37,39] that the OZ equation can be solved directly for the different components in real space. Results of equivalent precision can be reached from inversion relations in k-space[78].

Defining the number density,

$$n(\mathbf{r}) = \int d\boldsymbol{\omega}\,\rho(\mathbf{r},\boldsymbol{\omega}), \tag{39}$$

and the polarization density,

$$\mathbf{P}(\mathbf{r}) = \mu\int d\boldsymbol{\omega}\,\boldsymbol{\omega}\,\rho(\mathbf{r},\boldsymbol{\omega}), \tag{40}$$

and injecting the expression 37 of $c(\mathbf{r}_{12},\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)$ into the functional of eq. 33, it can be shown that external and excess terms can be written as functionals of $n(\mathbf{r})$ and $\mathbf{P}(\mathbf{r})$ instead of the much more complex variable $\rho(\mathbf{r},\boldsymbol{\omega})$, namely

$$\begin{aligned}\mathcal{F}_{ext}[n,\mathbf{P}] &= \int d\mathbf{r}\,\Phi_{LJ}(\mathbf{r}) - \int d\mathbf{r}\,\mathbf{P}(\mathbf{r})\cdot\mathbf{E}_q(\mathbf{r})\\ \mathcal{F}_{exc}[n,\mathbf{P}] &= -\frac{k_BT}{2}\int d\mathbf{r}_1 d\mathbf{r}_2\,c_S(r_{12})\Delta n(\mathbf{r}_1)\cdot\Delta n(\mathbf{r}_2)\\ &\quad -\frac{k_BT}{2\mu^2}\int d\mathbf{r}_1 d\mathbf{r}_2\,c_\Delta(r_{12})\mathbf{P}(\mathbf{r}_1)\cdot\mathbf{P}(\mathbf{r}_2)\\ &\quad -\frac{k_BT}{2\mu^2}\int d\mathbf{r}_1 d\mathbf{r}_2\,c_D(r_{12})\left[3(\mathbf{P}(\mathbf{r}_1)\cdot\hat{\mathbf{r}}_{12})\,(\mathbf{P}(\mathbf{r}_2)\cdot\hat{\mathbf{r}}_{12}) - \mathbf{P}(\mathbf{r}_1)\cdot\mathbf{P}(\mathbf{r}_2)\right].\end{aligned} \tag{41}$$

At this stage, the expression of the ideal term can be kept unchanged as a function of $\rho\,(\mathbf{r},\omega)$ as in eq. 33 and the minimization of the whole functional still performed with

respect to $\rho(\mathbf{r},\omega)$. The above expressions of the nonlocal excess free energy requires to perform FFT's for $n(\mathbf{r}), \mathbf{P}(\mathbf{r})$ rather than for $\rho(\mathbf{r},\omega)$ and this reduces considerably the computation time. We can go even a little bit further, and show that the ideal part itself can be expressed as a functional of $n(\mathbf{r})$ and $\mathbf{P}(\mathbf{r})$[37], namely

$$\begin{aligned}\mathcal{F}_{id}[n,\mathbf{P}] =& k_B T \int d\mathbf{r}\, \left(n(\mathbf{r}) \ln\left(\frac{n(\mathbf{r})}{n_0}\right) - n(\mathbf{r}) + n_0 \right) \\ &+ k_B T \int d\mathbf{r}\, n(\mathbf{r}) \left(\ln\left[\frac{\mathcal{L}^{-1}(\frac{P(\mathbf{r})}{\mu\, n(\mathbf{r})})}{\sinh(\mathcal{L}^{-1}(\frac{P(\mathbf{r})}{\mu\, n(\mathbf{r})}))} \right] + \frac{P(\mathbf{r})}{\mu\, n(\mathbf{r})}\, \mathcal{L}^{-1}(\frac{P(\mathbf{r})}{\mu\, n(\mathbf{r})}) \right).\end{aligned} \tag{42}$$

In the second, polarization term, $\mathcal{L}$ designates the Langevin function and $\mathcal{L}^{-1}$ its inverse; $P(\mathbf{r})$ is the modulus of the polarization vector $\mathbf{P}(\mathbf{r})$. The linearization of this term for small polarization fields yields the correct electrostatic limit, namely

$$\mathcal{F}_{id}[n,\mathbf{P}] = \int d\mathbf{r}\, \frac{\mathbf{P}(\mathbf{r})^2}{2\alpha_d n(\mathbf{r})}, \tag{43}$$

where $\alpha_d = \mu^2/3k_BT$ is the usual equivalent polarizability of a dipole μ at the temperature T. One recognizes the expression of the polarization free-energy in a medium with local electric susceptibility $\chi_e(\mathbf{r}) = \alpha_d n(\mathbf{r})$.

Although the functional is now complete and usable as such, we proceed by looking at an equivalent of eq. 41 involving susceptibilities rather than direct correlation functions. We introduce the longitudinal and transverse polarization in k-space

$$\begin{aligned}\mathbf{P}_L(\mathbf{k}) =& \quad (\mathbf{P}(\mathbf{k})\cdot\hat{\mathbf{k}})\,\hat{\mathbf{k}} \\ \mathbf{P}_T(\mathbf{k}) =& \quad \mathbf{P}(\mathbf{k}) - \mathbf{P}_L(\mathbf{k}),\end{aligned} \tag{44}$$

where $\hat{\mathbf{k}} = \mathbf{k}/k$. The electrostatic part of the excess free energy in eq.41 can be written in k-space

$$\begin{aligned}\mathcal{F}_{exc}^{elec} =& -\frac{1}{2}\frac{k_BT}{\mu^2}\int d\mathbf{k}\, c_\Delta(k)\, \mathbf{P}(\mathbf{k})\cdot\mathbf{P}(-\mathbf{k}) \\ &- \frac{1}{2}\frac{k_BT}{\mu^2}\int d\mathbf{k}\, c_D(k)\, \left[3(\mathbf{P}(\mathbf{k})\cdot\hat{\mathbf{k}})(\mathbf{P}(-\mathbf{k})\cdot\hat{\mathbf{k}}) - \mathbf{P}(\mathbf{k})\cdot\mathbf{P}(-\mathbf{k})\right].\end{aligned} \tag{45}$$

This can be easily rearranged into

$$\mathcal{F}_{exc}^{elec} = -\frac{1}{2}\frac{k_BT}{\mu^2}\left[d\mathbf{k}\, c_-(k)\, \mathbf{P}_T(\mathbf{k})\cdot\mathbf{P}_T(-\mathbf{k}) + \int d\mathbf{k}\, c_+(k)\, \mathbf{P}_L(\mathbf{k})\cdot\mathbf{P}_L(-\mathbf{k})\right] \tag{46}$$

with the usual definitions[10,79,80]:

$$c_-(k) = c_\Delta(k) - c_D(k) \tag{47}$$

$$c_+(k) = c_\Delta(k) + 2c_D(k). \tag{48}$$

We use now the relations between $c_-(k)$ and $c_+(k)$ and the longitudinal and transverse dielectric constant $\epsilon_L(k)$ and $\epsilon_T(k)$, or, alternatively, the longitudinal and transverse di-

electric susceptibilities $\chi_L(k)$ and $\chi_T(k)$ (see Refs[10,79–82])

$$1 - \frac{n_0}{3} c_+(k) = \frac{3y}{1 - 1/\epsilon_L(k)} = \frac{3y}{4\pi\chi_L(k)}, \tag{49}$$

$$1 - \frac{n_0}{3} c_-(k) = \frac{3y}{\epsilon_T(k) - 1} = \frac{3y}{4\pi\chi_T(k)}, \tag{50}$$

with $y = \mu^2 n_0 / 9 k_B T \epsilon_0$, such that

$$\begin{aligned} F_{exc}^{elec} = & - \frac{3k_BT}{n_0\mu^2} \int d\mathbf{k}\, \mathbf{P}(\mathbf{k}) \cdot \mathbf{P}(-\mathbf{k}) \\ & + \frac{1}{8\pi\epsilon_0} \left[\int d\mathbf{k}\, \frac{\mathbf{P}_T(\mathbf{k}) \cdot \mathbf{P}_T(-\mathbf{k})}{\chi_T(k)} + \int d\mathbf{k}\, \frac{\mathbf{P}_L(\mathbf{k}) \cdot \mathbf{P}_L(-\mathbf{k})}{\chi_L(k)} \right]. \end{aligned} \tag{51}$$

(beware of the definition of χ_L, with or without a 4π factor[79,81]).

At the end, one can gather all the above equations, including eqs. 13,41,51 to get the following functional for a dipolar fluid, defined in terms of the density and dielectric susceptibilities

$$\begin{aligned} \mathcal{F}[n, \mathbf{P}] = & k_BT \int d\mathbf{r} d\boldsymbol{\omega} \left[\rho(\mathbf{r}, \boldsymbol{\omega}) \ln \left(\frac{4\pi\rho(\mathbf{r}, \boldsymbol{\omega})}{n_0} \right) - \rho(\mathbf{r}, \boldsymbol{\omega}) + \frac{n_0}{4\pi} \right] \\ & + \int d\mathbf{r}\, \Phi_{LJ}(\mathbf{r})\, n(\mathbf{r}) - \frac{k_BT}{2n_0} \int d\mathbf{r}\, \Delta n(\mathbf{r})^2 \\ & + \frac{k_BT}{2} \int d\mathbf{r}_1 d\mathbf{r}_2\, \chi_n^{-1}(r_{12})\, \Delta\rho(\mathbf{r}_1) \Delta\rho(\mathbf{r}_2) \\ & - \int d\mathbf{r}\, \mathbf{P}(\mathbf{r}) \cdot \mathbf{E}_q(\mathbf{r}) - \frac{3k_BT}{n_0\mu^2} \int d\mathbf{r}\, \mathbf{P}(\mathbf{r})^2 \\ & + \frac{1}{8\pi\epsilon_0} \int d\mathbf{r}_1 d\mathbf{r}_2\, \chi_T^{-1}(r_{12})\, \mathbf{P}_T(\mathbf{r}_1) \cdot \mathbf{P}_T(\mathbf{r}_2) \\ & + \frac{1}{8\pi\epsilon_0} \int d\mathbf{r}_1 d\mathbf{r}_2\, \chi_L^{-1}(r_{12})\, \mathbf{P}_L(\mathbf{r}_1) \cdot \mathbf{P}_L(\mathbf{r}_2) + \mathcal{F}_B[n, \mathbf{P}]. \end{aligned} \tag{52}$$

The ideal part can also be taken as in eq. 42 so that the whole functional can be minimized with respect to $n(\mathbf{r})$ and $\mathbf{P}(\mathbf{r})$. The bridge term can be neglected (HNC approximation) or approximated as a functional of $n(\mathbf{r})$ only, $\mathcal{F}_B[n]$, for example using the hard-sphere bridge functional of Section 2.2

As a short illustration, Fig. 4a shows the accuracy of the MDFT approach (within the HNC approximation) for the microscopic structure of the Stockmayer solvent around neutral and charged spherical solutes[37,38,63]. The MDFT results are compared to direct MD simulations of the solute embedded in the solvent. They do appear very satisfactory and account accurately for the shape of the peaks and their variation with charge and size (despite a slight overestimation of the first peak height for the neutral solute). Fig. 4b illustrates the case of a multisite polar molecule (here a three-site model of the acetonitrile molecule) with similar conclusions. An application to a more complex molecular system, namely the three-dimensional solvation structure close to an atomistically resolved clay, is illustrated in Fig. 5 and described further in Ref.[66].

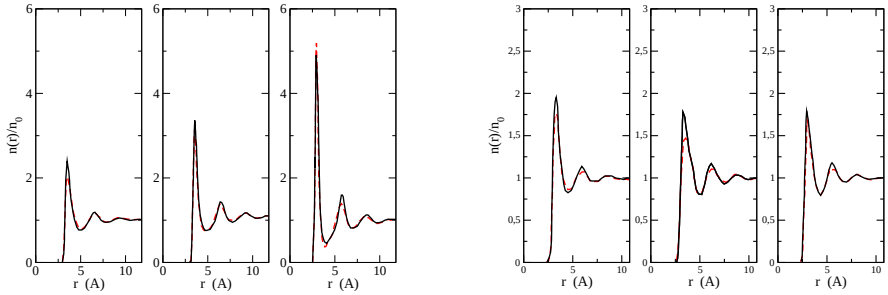


Figure 4: Left: Reduced density of the Stockmayer solvent around various solutes. MDFT results (solid black lines) are compared to MD simulation results (dashed red lines). From left to right: CH_4, Cl^-, K^+. Right: Same than for the various sites of an acetonitrile molecule dissolved in the Stockmayer solvent. From left to right: CH_3, C, N.

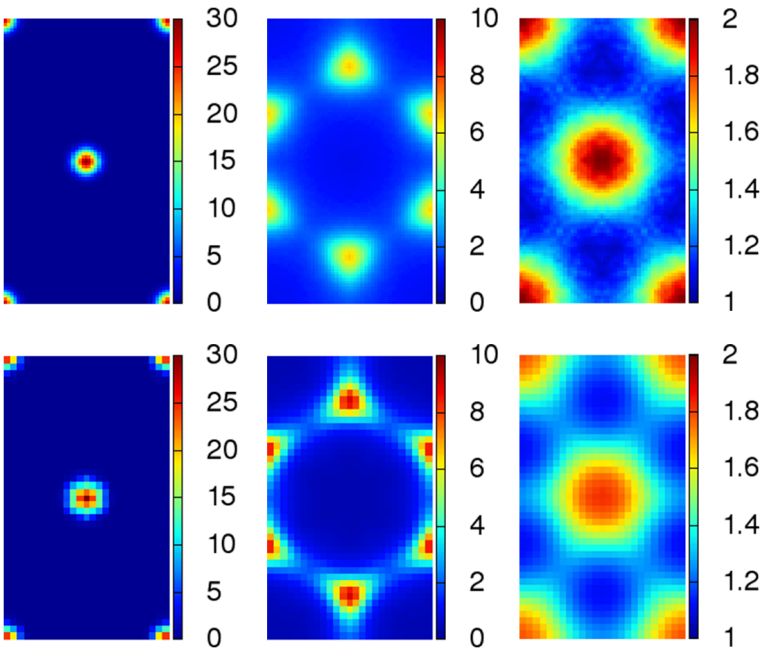


Figure 5: Two-dimensional maps of the solvent number density $n(\mathbf{r})/n_0$ in three different planes close to a neutral clay surface, as calculated by molecular dynamics (top) and HRF-MDFT (bottom). Those planes correspond to a prepeak (left), the first maximum (center) and second maximum (right) of the out-of plane mean solvent density. See Ref.[66]

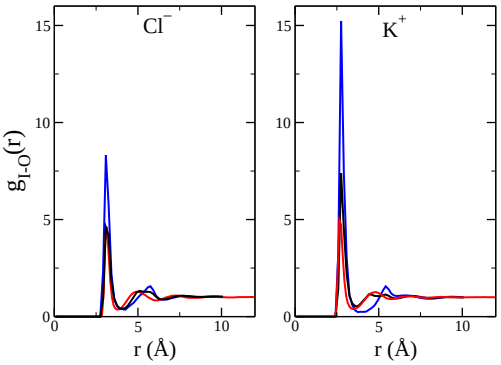


Figure 6: Ion-oxygen pair distribution functions for chloride and potassium ions in SPC/E water computed by MD (red lines) or MDFT without (blue lines) or with the three body term described in Ref.[68] (black lines)

3.3 Extension to water and arbitrary molecular solvents

Water is a special case, certainly by its very subtle physics, but also for the fact that the most popular molecular models fall in the category of simple point charge models with a single Lennard-Jones center (usually centered on the oxygen atom) and distributed point charges. In that case, it was shown recently that the functional just displayed in eq. 52 (with the linear orientation vector $\boldsymbol{\omega}$ substituted by the three-angle orientation ω, and 4π by $8\pi^2$) is perfectly applicable if the dipolar polarization $\mathbf{P}(\mathbf{r})$ is replaced by a multipolar polarization vector, accounting for the full charge distribution of the water molecule, and defined in k-space by

$$\mathbf{P}(\mathbf{k}) = \int d\omega \boldsymbol{\mu}(\mathbf{k},\omega)\rho(\mathbf{k},\omega) \tag{53}$$

with

$$\boldsymbol{\mu}(\mathbf{k},\omega) = -i \sum_m q_m \frac{\mathbf{s}_m(\omega)}{\mathbf{k}\cdot\mathbf{s}_m(\omega)} \left(e^{i\,\mathbf{k}\cdot\mathbf{s}_m(\omega)} - 1\right) \tag{54}$$

$$= \boldsymbol{\mu}(\omega) + \frac{i}{2}\sum_m q_m \left(\mathbf{k}\cdot\mathbf{s}_m(\omega)\right)\mathbf{s}_m(\omega) + ..., \tag{55}$$

being defined as the polarization, of a single molecule located at the origin. $\mathbf{s}_m(\Omega)$ designates the location of the m^{th} atomic site for a given orientation ω. It reduces to the usual molecular dipole $\boldsymbol{\mu}(\omega) = \sum_m q_m\,\mathbf{s}_m(\omega)$ at dominant order in $\mathbf{k}$. The multipolar dielectric susceptibilities $\chi_L(k)$ and $\chi_T(k)$ entering in eq. 52 can be either computed from MD simulations of the pure liquid, according to the procedure in Refs[81,82], or inferred from experiments.

Although giving already sensible results for rather complex systems[69], it was shown in that the HNC approximation $\mathcal{F}_B = 0$ turns unfortunately short for describing the solvation of hydrophobes[65,67], as well as that of molecular solutes giving raise to strong H-bonds[68].

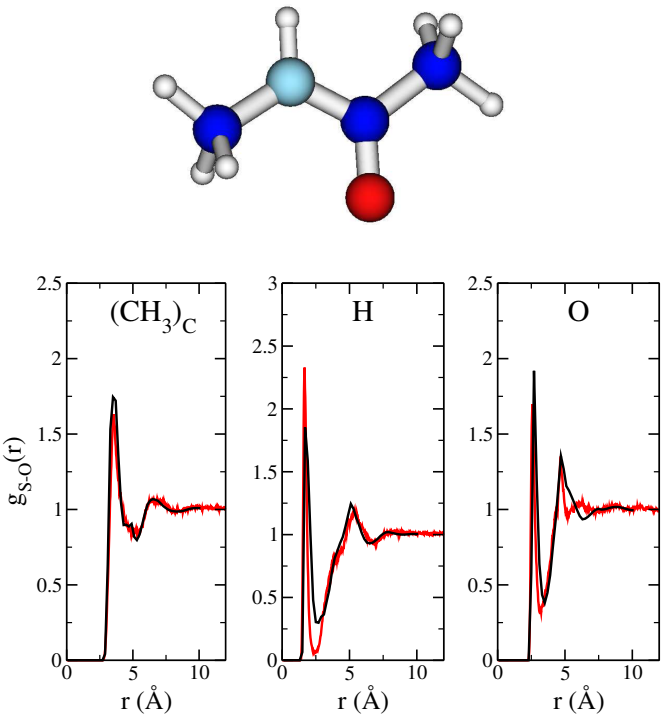


Figure 7: A selection of solute site-oxygen pair distribution functions for the n-methyl-acetamide molecule $CH_3NHCOCH_3$ (shown on top) computed by MD (red lines) or MDFT (black lines), including the three-body term described in Ref.[68].

Three-body corrections, including a spherical HS bridge, or a three-body term re-enforcing tetrahedral order, have to be added to give correct solvation structure and thermodynamics. This is illustrated in Fig. 6 for the hydration structure around monovalent ions. Fig. 7 shows the water structure obtained by MDFT around a N-methyl-acetamide molecule (the prototype for a NH-CO peptide motif), including the three-body correction term in the functional.[68]

For a general solvent with more complex geometry, and described by more than one Lennard-Jones center, the full angular-dependent functional of Sec. 3.1 has to be adopted, and the necessary input remains the full angular-dependent direct correlation function $c(\mathbf{r}_{12}, \boldsymbol{\omega}_1, \boldsymbol{\omega}_2)$. Remaining in the HNC approximation, this formulation was applied with some success to the study of charge transfer processes in acetonitrile[64].

The MDFT approach is still under current development, as are related site-DFT approaches[71], for practical applications such as the systematic prediction of solvation free energies[70]. Classical density functional theories are expected to provide soon an alternative to the 3D-RISM approach, which is nowadays becoming quite popular for applications in biological and material sciences –despite some intrinsic theoretical limitations that specialists are aware of.

References

1. G. M. Torrie and J. P. Valleau, *Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling*, Journal of Computational Physics, **23**, no. 2, 187–199, 1977.
2. J. Valleau, in: Classical and Quantum Dynamics in Condensed Phase Simulations, B. J. Berne, G. Ciccotti, and D.F. Cocker, (Eds.), p. 97, World Scientific Co. 1998.
3. G. Ciccotti, in: Classical and Quantum Dynamics in Condensed Phase Simulations, B. J. Berne, G. Ciccotti, and D.F. Cocker, (Eds.), p. 159, World Scientific Co. 1998.
4. Benoit Roux and Thomas Simonson, *Implicit solvent models*, Biophysical Chemistry, **78**, no. 1–2, 1–20, 1999.
5. Barry Honig, Kim Sharp, and An Suei Yang, *Macroscopic models of aqueous solutions: biological and chemical applications*, The Journal of Physical Chemistry, **97**, no. 6, 1101–1109, 1993.
6. Nathan A. Baker, David Sept, Simpson Joseph, Michael J. Holst, and J. Andrew McCammon, *Electrostatics of nanosystems: Application to microtubules and the ribosome*, Proceedings of the National Academy of Sciences, **98**, no. 18, 10037–10041, 2001.
7. M. Marchi, D. Borgis, N. Lévy, and P. Ballone, *A dielectric continuum molecular dynamics method*, J. Chem. Phys., **114**, 4377–4385, 2001.
8. L. T. Cheng, J. Dzubiella, and J. A. McCammon, *Application of the level-set method to the implicit solvation of nonpolar molecules*, J. Chem. Phys., **127**, 084503, 2007.
9. L. T. Ti Cheng, Y. Xie, J. Dzubiella, J. A. McCammon, J. Che, and Bo Li, *Coupling the level-set method with molecular mechanics for variational implicit solvation of nonpolar molecules*, J. Chem. Theor. Comp., **5**, 257–266, 2009.
10. J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, Academic Press, London, 1989.
11. C. G. Gray and K. E. Gubbins, *Theory of Molecular Fluids: I: Fundamentals*, OUP Oxford, 1984.
12. Christopher G. Gray, Keith E. Gubbins, and Christopher G. Joslin, *Theory of Molecular Fluids: Volume 2: Applications*, OUP Oxford, 2011.
13. D. Chandler and H.C. Hendersen, *Optimized Cluster Expansions for Classical fluids - theory of molecular liquids*, J. Chem. Phys., **57**, 1930, 1972.
14. F. Hirata and P. J. Rossky, *An extended rism equation for molecular polar fluids*, Chem. Phys. Lett., **83**, 329, 1981.
15. F. Hirata, B. M. Pettitt, and P. J. Rossky, *Application of an extended rism equation to dipolar and quadrupolar fluids*, J. Chem. Phys., **77**, 509, 1982.
16. G. Reddy, C. P. Lawrence, J. L. Skinner, and A. Yethiraj, *Liquid state theories for the structure of water*, J. Chem. Phys., **119**, 13012, 2003.
17. L. Blum and A. J. Torruella, *Invariant expansion for 2-body correlations - thermodynamic functions, scattering, and Ornstein-Zernike equation*, J. Chem. Phys., **56**, 303, 1972.
18. L. Blum, *Invariant expansion - Ornstein-Zernike equation for nonspherical molecules and an extended solution to mean spherical model*, J. Chem. Phys., **57**, 1862, 1972.
19. G. N. Patey, *Integral-equation theory for dense dipolar hard-sphere fluid*, Mol. Phys., **34**, 427, 1977.

20. S. L. Carnie and G. N. Patey, *Fluids of polarizable hard-spheres with dipoles and tetrahedral quadrupoles - integral-equation results with application to liquid water*, Mol. Phys., **47**, 1129, 1982.
21. P. H. Fries and G. N. Patey, *The solution of the hypernetted-chain approximation for fluids of nonspherical particles - a general-method with application to dipolar hard-spheres*, J. Chem. Phys., **82**, 429, 1985.
22. J. Richardi, P. H. Fries, and H. Krienke, *The solvation of ions in acetonitrile and acetone: A molecular Ornstein-Zernike study*, J. Chem. Phys., **108**, 4079, 1998.
23. J. Richardi, C. Millot, and P. H. Fries, *A molecular Ornstein-Zernike study of popular models for water and methanol*, J. Chem. Phys., **110**, 1138, 1999.
24. K. M. Dyer, J. S. Perkyns, and B. M. Pettitt, *A site-renormalized molecular fluid theory*, J. Chem. Phys., **127**, 194506, 2007.
25. K. M. Dyer, J. S. Perkyns, G. Stell, and B. M. Pettitt, *A molecular site-site integral equation that yields the dielectric constant*, J. Chem. Phys., **129**, 104512, 2008.
26. D. Chandler, *Gaussian field model of fluids with an application to polymeric fluids*, Phys. Rev. E, **48**, 2898, 1993.
27. P. Rein ten Wolde, S. X. Sun, and D. Chandler, *Model of a fluid at small and large length scales and the hydrophobic effect*, Phys. Rev. E, **65**, 011201, 2002.
28. R. Evans, *The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids*, Advances in Physics, **28**, no. 2, 143, 1979.
29. R. Evans, *Fundamentals of Inhomogeneous Fluids*, Marcel Dekker, Incorporated, 1992.
30. R. Evans, "Density functional theory for inhomogeneous fluids: Simple fluids in equilibrium", in: Lecture notes at 3rd Warsaw School of Statistical Physics. June 2009.
31. David Chandler, John D. McCoy, and Sherwin J. Singer, *Density functional theory of nonuniform polyatomic systems. I. General formulation*, The Journal of Chemical Physics, **85**, no. 10, 5971, 1986.
32. David Chandler, John D. McCoy, and Sherwin J. Singer, *Density functional theory of nonuniform polyatomic systems. II. Rational closures for integral equations*, The Journal of Chemical Physics, **85**, no. 10, 5977, 1986.
33. T. Biben, J. P. Hansen, and Y. Rosenfeld, *Generic density functional for electric double layers in molecular solvent*, Phys. Rev. E, **57**, R3727–3730, 1998.
34. Anna Oleksy and Jean-Pierre Hansen, *Microscopic density functional theory of wetting and drying of a solid substrate by an explicit solvent model of ionic solutions*, Molecular Physics, **107**, no. 23-24, 2609–2624, 2009.
35. Anna Oleksy and Jean-Pierre Hansen, *Wetting of a solid substrate by a "civilized" model of ionic solutions*, The Journal of Chemical Physics, **132**, no. 20, 204702–204702–13, 2010.
36. Anna Oleksy and Jean-Pierre Hansen, *Wetting and drying scenarios of ionic solutions*, Molecular Physics, **109**, no. 7-10, 1275–1288, 2011.
37. R. Ramirez, R. Gebauer, M. Mareschal, and D. Borgis, *Density functional theory of solvation in a polar solvent: Extracting the functional from homogeneous solvent simulations*, Phys. Rev. E, **66**, 306, 2002.
38. R. Ramirez and D. Borgis, *Density functional theory of solvation and its relation to implicit solvent models*, J. Phys. Chem. B, **109**, 6754, 2005.
39. R. Ramirez, M. Mareschal, and D. Borgis, *Direct correlation functions and the density functional theory of polar solvents*, Chem. Phys., **319**, 261, 2005.

40. R. D. Coalson, A. M. Walsh, A. Duncan, and N. Ben-Tal, *Statistical mechanics of a Coulomb gas with finite size particles: A lattice fluid theory approach*, J. Chem. Phys., **102**, 4584, 1995.
41. R. D. Coalson and A. Duncan, *Statistical Mechanics of a Multipolar Gas: A Lattice Field Theory Approach*, J. Phys. Chem. B, **100**, 2612, 1996.
42. R. Coalson and T. Beck, *Encyclopedia of Computational Chemistry*, vol. 3, Wiley, New York, 1998.
43. C. Azuara, E. Lindahl, and P. Koehl, *PDB_Hydro: incorporating dipolar solvents with variable density in the Poisson-Boltzmann treatment of macromolecule electrostatics*, Nucleic Ac. Res., **34**, 38, 2006.
44. C. Azuara, H. Orland, M. Bon, P. Koehl, and M. Delarue, *Incorporating dipolar solvents with variable density in Poisson-Boltzmann electrostatics*, Biophys. J., **95**, 5587, 2008.
45. D. Beglov and B. Roux, *An Integral Equation to Describe the Solvation of Polar Molecules in Liquid Water*, J. Phys. Chem. B, **101**, 7821, 1997.
46. A. Kovalenko and F. Hirata, *Three-dimensional density profiles of water in contact with a solute of arbitrary shape; a RISM approach*, Chem. Phys. Lett., **290**, 237, 1998.
47. Ed. F. Hirata, *Molecular Theory of Solvation*, Kluwer Academic Publishers, Dordrecht, 2003.
48. N. Yoshida, T. Imai, S. Phongphanphanee, A. Kovalenko, and F. Hirata., *Molecular Recognition in Biomolecules Studied by Statistical-Mechanical Integral-Equation Theory of Liquids*, J. Phys. Chem. B, **113**, 873–886, 2009.
49. Volodymyr P. Sergiievskyi and Maxim V. Fedorov, *3DRISM Multigrid Algorithm for Fast Solvation Free Energy Calculations*, Journal of Chemical Theory and Computation, **8**, no. 6, 2062–2070, 2012.
50. David S. Palmer, Volodymyr P. Sergiievskyi, Frank Jensen, and Maxim V. Fedorov, *Accurate calculations of the hydration free energies of druglike molecules using the reference interaction site model*, The Journal of Chemical Physics, **133**, no. 4, 044104, 2010.
51. H Lowen, *Density functional theory of inhomogeneous classical fluids: recent developments and new perspectives*, Journal of Physics: Condensed Matter, **14**, no. 46, 11897–11905, 2002.
52. Yaakov Rosenfeld, *Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing*, Physical Review Letters, **63**, no. 9, 980–983, 1989.
53. E. Kierlik and M. L. Rosinberg, *Free-energy density functional for the inhomogeneous hard-sphere fluid: Application to interfacial adsorption*, Physical Review A, **42**, no. 6, 3382–3387, 1990.
54. E. Kierlik and M. L. Rosinberg, *Density-functional theory for inhomogeneous fluids: Adsorption of binary mixtures*, Physical Review A, **44**, no. 8, 5025–5037, 1991.
55. R. Roth, R. Evans, A. Lang, and G. Kahl, J. Phys. : Condens. Matter, **14**, 12063, 2002.
56. Yang-Xin Yu and Jianzhong Wu, *Structures of hard-sphere fluids from a modified fundamental-measure theory*, The Journal of Chemical Physics, **117**, no. 22, 10156, 2002.

57. R. Roth, *Fundamental measure theory for hard-sphere mixtures: A review*, J. Phys.: Condens. Matter, **22**, 063102, 2010.
58. J. Wu and Z. Li, *Density functional theory for complex fluids*, Ann. Rev. Phys. Chem., **58**, 85, 2007.
59. J. Wu, in: Molecular Thermodynamics of Complex Systems, X. Lu and Y. Hu, (Eds.), Springer. 2009.
60. P. I. Texeira and M. M. Telo da Gama, *Density-functional theory for the interfacial properties of a dipolar fluid*, J. Phys.: Condens. Matter, **3**, 111–125, 1991.
61. P. Frodl and S. Dietrich, *Bulk and interfacial properties of polar and molecular fluids*, Phys. Rev. A, **45**, 7330, 1992.
62. L. Gendre, R. Ramirez, and D. Borgis, *Classical density functional theory of solvation in molecular solvents: Angular grid implementation*, Chem. Phys. Lett., **474**, 366, 2009.
63. S. Zhao, R. Ramirez, R. Vuilleumier, and D. Borgis, *Molecular density functional theory of solvation: From polar solvents to water*, J. Chem. Phys., **134**, 194102, 2011.
64. D. Borgis, D. Gendre, and R. Ramirez, *Molecular Density Functional Theory: Application to Solvation and Electron-Transfer Thermodynamics in Polar Solvents*, J. Phys. Chem. B, **116**, 2012.
65. Maximilien Levesque, Rodolphe Vuilleumier, and Daniel Borgis, *Scalar fundamental measure theory for hard spheres in three dimensions: Application to hydrophobic solvation*, The Journal of Chemical Physics, **137**, no. 3, 034115–1–034115–9, 2012.
66. Maximilien Levesque, Virginie Marry, Benjamin Rotenberg, Guillaume Jeanmairet, Rodolphe Vuilleumier, and Daniel Borgis, *Solvation of complex surfaces via molecular density functional theory*, The Journal of Chemical Physics, **137**, no. 22, 224107–224107–8, 2012.
67. Guillaume Jeanmairet, Maximilien Levesque, and Daniel Borgis, *Molecular density functional theory of water describing hydrophobicity at short and long length scales*, The Journal of Chemical Physics, **139**, no. 15, 154101–1–154101–9, 2013.
68. Guillaume Jeanmairet, Maximilien Levesque, Rodolphe Vuilleumier, and Daniel Borgis, *Molecular Density Functional Theory of Water*, The Journal of Physical Chemistry Letters, **4**, 619–624, 2013.
69. Guillaume Jeanmairet, Virginie Marry, Maximilien Levesque, Benjamin Rotenberg, and Daniel Borgis, *Hydration of clays at the molecular scale: the promising perspective of classical density functional theory*, Molecular Physics, **112**, no. 9-10, 1320–1329, 2014.
70. Volodymyr P. Sergiievskyi, Guillaume Jeanmairet, Maximilien Levesque, and Daniel Borgis, *Fast Computation of Solvation Free Energies with Molecular Density Functional Theory: Thermodynamic-Ensemble Partial Molar Volume Corrections*, The Journal of Physical Chemistry Letters, **5**, no. 11, 1935–1942, 2014.
71. Yu Liu, Shuangliang Zhao, and Jianzhong Wu, *A Site Density Functional Theory for Water: Application to Solvation of Amino Acid Side Chains*, Journal of Chemical Theory and Computation, **9**, no. 4, 1896–1908, 2013.
72. J. P. Hansen, in: The Physics and Chemistry of Aqueous Ionic Solutions, M. C. Bellissent-Funel and G. W. Neilson, (Eds.), Kluwer Academic Publishers, Dortrecht, Holland, 1987.

73. S. Phan, E. Kierlik, M. L. Rosinberg, B. Bildstein, and G. Kahl, *Equivalence of two free-energy models for the inhomogeneous hard-sphere fluid*, Physical Review E, **48**, no. 1, 618–620, 1993.
74. Yaakov Rosenfeld, *Free energy model for inhomogeneous fluid mixtures: Yukawa-charged hard spheres, general interactions, and plasmas*, The Journal of Chemical Physics, **98**, no. 10, 8126–8148, 1993.
75. M. Oettel, *Integral equations for simple fluids in a general reference functional approach*, Journal of Physics: Condensed Matter, **17**, no. 3, 429, 2005.
76. Y. Tang, *First-order mean spherical approximation for inhomogeneous fluids*, J. Chem. Phys., **121**, 10605–10610, 2004.
77. T. Lazaridis, *Inhomogeneous Fluid Approach to Solvation Thermodynamics. 2. Applications to Simple Fluids*, J. Phys. Chem. B, **102**, 3542–3550, 1998.
78. L. Gendre, *Density functional theory of molecular liquids: Application to solvation in polar solvents*, PhD thesis, Université d'Evry-Val-d'Essonne, Evry, France, 2008.
79. F. O. Raineri, H. Resat, and H. L. Friedman, *Static longitudinal dielectric function of model molecular fluids*, J. Chem. Phys., **96**, 3068, 1992.
80. F. O. Raineri and H. L. Friedman, *Static transverse dielectric function of model molecular fluids*, J. Chem. Phys., **98**, 8910, 1993.
81. P. A. Bopp, A. A. Kornyshev, and G. Sutmann, *Static Nonlocal Dielectric Function of Liquid Water*, Phys. Rev. Lett., **76**, 1281, 1996.
82. P. A. Bopp, A. A. Kornyshev, and G. Sutmann, *Frequency and wave-vector dependent dielectric function of water: Collective modes and relaxation spectra*, J. Chem. Phys., **109**, 1939, 1998.

Accurate Coarse-Grained Potentials for Soft Matter Systems

Ronald Blaak, Barbara Capone, Christos N. Likos, and Lorenzo Rovigatti

Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria

E-mail: {ronald.blaak, barbara.capone,christos.likos, lorenzo.rovigatti} @univie.ac.at

In this Chapter, we present the general statistical-mechanical theory for the derivation of effective Hamiltonians for arbitrary many-body systems in thermodynamic equilibrium. The approach towards a simplified, yet accurate, *coarse-grained* view of a many-body system has proven to be particularly fruitful in the realm of the Physics of Complex Fluids, a synonym for the broad research area of Soft Matter Science, which encompasses aspects of Physics, Chemistry, Materials Science and Bioscience. After the exposition of the general formalism, we present the methods by which one can derive, in a manageable and accurate way, effective Hamiltonians for a veriety of physical systems, by exposing mutually complementary theoretical and computations techniques to this end. Further, we provide extensive exposures of specific and diverse examples from current research in soft matter systems, for which the coarse-graining approach has been applied with remarkable success.

1 Introduction

Soft matter is a subfield of condensed matter that deals with materials whose dominant physical behavior occurs on an energy scale of the order of thermal energy. Soft matter systems are extremely complex materials, typically composed of mesoscopic particles i.e., particles with sizes between the nanometer and the micrometer scale, dispersed into a solvent constituted by much smaller molecules (typically of atomic dimensions). These *colloidal suspensions* can, in turn, be complemented by co-solutes such as polymers, ions, etc. The huge difference in the length scales associated to the different species in solution results in a very large range of characteristic time scales, which span many orders of magnitude; the dynamics of the solvent molecules, which are very small and thus very fast, needs to be resolved on the time scale of ps, while the typical characteristic diffusion times of colloidal particles is of the order of μs.[1] On one hand, this makes it very impractical to keep track of the solvent degrees of freedom, numerically as well as theoretically. On the other hand, in the majority of the cases we are only interested in the structure and dynamics of the solute. We can thus think of an effective *one-component* description of the system, where the effect of the solvent can be taken into account indirectly by regarding it as a "mediator" of the colloid-colloid interaction.

The process of going from a full-detailed description to a simpler one, also known as *coarse-graining*, can be generalised to any kind of system, and can also be iterated to further reduce the complexity. For instance, polymer systems are good examples for the application of a hierarchical coarse-graining procedure: at the first level, the solvent molecules can be treated implicitly and the polymer monomers can be identified by their centres of mass only, obtaining a monomer-resolved description.[2–4] The resolution, and thus the number of degrees of freedom, can be further lowered by considering polymeric chains as composed by several blobs, each containing a certain number of monomers.[2,5]

Finally, one can go as far as describing a chain with its center of mass, so as to map each chain to a single soft sphere.[6]

The two main strategies used to perform a coarse-graining are the *top-down* and *bottom-up* approaches. With the former, the matching between the original and coarse-grained systems is carried out in a heuristic way by mapping mesoscopic or macroscopic quantities, such as the persistence length, the melting temperature or the elastic modulus. Examples of top–down coarse-grained models can be found in Refs.[7,8] By contrast, *bottom-up* approaches start from the microscopic description of the system, usually given in the form of the full Hamiltonian, and rigorously trace out some of the microscopic degrees of freedom.[1]

In this Chapter we first provide a theoretical basis for the bottom-up approach, and then describe in depth a few reliable computational methods that can be used to apply this concept to real-world soft-matter systems. The rest of this work is organized as follows: the general theory concerning the definition of the effective Hamiltonian and the structure of the resulting effective interactions is presented in Section 2. On the basis of the formalism and the properties of this effective interaction, we present in Section 3 the methods by which it can be calculated or measured by means of theoretical or computational techniques. In the following three sections, we present applications of the general formalism to specific soft-matter systems: generic mixtures of size-asymmetric, hard-soft colloidal particles are discussed in Section 4. Dendrimers of various generations, both neutral and charged, are discussed in Section 5, whereas in Section 6 we focus our attention on ring polymers. Finally, in section 7, we summarize and draw our conclusions.

2 The Effective Hamiltonian

2.1 General considerations and formal definitions

Consider a ν-component ($\nu \geq 1$) classical many-body system. There exist N_α particles of species α enclosed in the macroscopic volume V and the system is at thermodynamic equilibrium at absolute temperature T. The quantities $\rho_\alpha \equiv N_\alpha/V$ are the corresponding partial densities, whereas $\rho = \sum_{\alpha=1}^{\nu} \rho_\alpha$ is the total density of the mixture. Each particle of species α has mass m_α, and the corresponding thermal de Broglie wavelength is $\Lambda_\alpha = \hbar\sqrt{2\pi\beta/m_\alpha}$, where $\beta = (k_\mathrm{B}T)^{-1}$, and k_B is Boltzmann's constant. Let, in addition, $\{\mathbf{s}_{j\alpha}\}$, $j = 1, 2, \ldots, N_\alpha$, denote the coordinates of the particles of species α and $\{\mathbf{p}_{j\alpha}\}$ be the corresponding momenta. The total Hamiltonian of the system, $\mathcal{H}$, can be written as the sum of the kinetic and interaction terms, $\mathcal{H} = K + U$. The kinetic energy is, of course, given as

$$K = \sum_{\alpha=1}^{\nu} \sum_{j=1}^{N_\alpha} \frac{\mathbf{p}_{j\alpha}^2}{2m_\alpha}, \tag{1}$$

and we will not specify any further the form of the potential energy function U apart from reminding the reader that it depends on the set of all coordinates of all species. Let us also introduce a shorthand for the configurational integral over the coordinates of species α acting on any quantity $\mathcal{O}$ that depends on the coordinates:

$$\mathrm{Tr}_\alpha \mathcal{O} \equiv \int \mathrm{d}\mathbf{s}_{1\alpha} \cdots \int \mathrm{d}\mathbf{s}_{N_\alpha\alpha} \mathcal{O}. \tag{2}$$

With these definitions, the canonical partition function $\mathcal{Z}$ of the system takes the form

$$\mathcal{Z} = \prod_{\alpha=1}^{\nu} \frac{1}{N_\alpha! \Lambda_\alpha^{3N_\alpha}} \mathrm{Tr}_\alpha \exp\left(-\beta U\right). \tag{3}$$

Suppose now we wish to calculate the expectation value of some quantity $\mathcal{Q}_\gamma$ whose instantaneous value depends solely on a subset $\{\mathbf{s}_{1\gamma}, \ldots, \mathbf{s}_{k\gamma}\}$, $k \leq N_\gamma$, of the coordinates of particles of the γ-species. This can be obtained easily as

$$\langle \mathcal{Q}_\gamma \rangle = \frac{1}{\mathcal{Z}} \prod_{\alpha=1}^{\nu} \frac{1}{N_\alpha! \Lambda_\alpha^{3N_\alpha}} \mathrm{Tr}_\alpha \left[\mathcal{Q}_\gamma \exp\left(-\beta U\right)\right]. \tag{4}$$

Moreover, we can go one step further and formally fix the variables to some prescribed values, $\mathbf{s}_{i\gamma} = \mathbf{S}_i$, $i = 1, 2, \ldots, k$. In a similar fashion, we can then calculate a generalized k-body 'connected correlation function' $G_\gamma^{(k)}(\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_k)$ between particles of species γ as:

$$G_\gamma^{(k)}(\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_k) = \frac{1}{\mathcal{Z}} \prod_{\alpha=1}^{\nu} \frac{1}{N_\alpha! \Lambda_\alpha^{3N_\alpha}} \mathrm{Tr}_\alpha \left[\exp\left(-\beta U\right) \prod_{i=1}^{k} \delta\left(\mathbf{s}_{i\gamma} - \mathbf{S}_i\right)\right]. \tag{5}$$

The physical idea behind coarse-graining is, as mentioned at the Introduction, to go over from a full Hamiltonian that contains the coordinates of all particles and all species as degrees of freedom, to one that features only those of a selected species, say γ. In doing so, we wish to maintain exactly both the thermodynamics of the full mixture, i.e., its partition function $\mathcal{Z}$, as well as all the expectation values of quantities involving the coordinates of species γ and only of γ. Formally, this is achieved by defining the effective interaction $\tilde{U}_\gamma(\{\mathbf{s}_{i\gamma}\})$ between particles of species γ through a partial trace in Eq. (3), in which the trace is carried out over all other species *excluding* $\alpha = \gamma$, i.e.:

$$\exp\left(-\beta \tilde{U}_\gamma\right) = \prod_{\alpha \neq \gamma}^{\nu} \frac{1}{N_\alpha! \Lambda_\alpha^{3N_\alpha}} \mathrm{Tr}_\alpha \exp\left(-\beta U\right). \tag{6}$$

From Eqs. (3) and (6) it is evident that with this definition the partition function remains invariant under the elimination of the other species and it can be equally well expressed in the effective one-component Hamiltonian as:

$$\mathcal{Z} = \frac{1}{N_\gamma! \Lambda_\gamma^{3N_\gamma}} \mathrm{Tr}_\gamma \exp\left(-\beta \tilde{U}_\gamma\right). \tag{7}$$

From the above definitions, it immediately follows that also the quantities of Eqs. (4) and (5) remain invariant if expressed now in the effective one-component description, i.e., we can write

$$\langle \mathcal{Q}_\gamma \rangle = \frac{1}{\mathcal{Z}} \frac{1}{N_\gamma! \Lambda_\gamma^{3N_\gamma}} \mathrm{Tr}_\gamma \left[\mathcal{Q}_\gamma \exp\left(-\beta \tilde{U}_\gamma\right)\right] \tag{8}$$

and

$$G_\gamma^{(k)}(\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_k) = \frac{1}{\mathcal{Z}} \frac{1}{N_\gamma! \Lambda_\gamma^{3N_\gamma}} \mathrm{Tr}_\gamma \left[\exp\left(-\beta \tilde{U}_\gamma\right) \prod_{i=1}^{k} \delta\left(\mathbf{s}_{i\gamma} - \mathbf{S}_i\right)\right]. \tag{9}$$

Formally, therefore, we have found a way to reduce the multicomponent system to a single-component one, which is described the the effective Hamiltonian $\mathcal{H}_{\rm eff}$ that features exclusively the coordinates and momenta of the species γ on which we are mainly interested, namely:

$$\mathcal{H}_{\rm eff} = \mathcal{H}_0 + \sum_{i=1}^{N_\gamma} \frac{\mathbf{p}_{i\gamma}^2}{2m_\gamma} + U_{\rm eff}(\{\mathbf{s}_{i\gamma}\}), \tag{10}$$

where we have written $\tilde{U}_\gamma = \mathcal{H}_0 + U_{\rm eff}$, separating a term $\mathcal{H}_0$ that does not depend on the coordinates $\{\mathbf{s}_{i\gamma}\}$, whose physical meaning and origing will be discussed later on.

What, if anything at all, have we achieved? Up to now, not much it seems, apart from writing down Eq. (6) and its immediate consequences. Indeed, from the practical point of view, one might say that exactly *nothing* has been achieved, for if we could carry out the integrations of Eq. (6), we might then as well carry out the remaining ones in Eq. (7) and obtain the partition function of the system. In such a case, the effective interaction would be just a trivial and unnecessary intermediate result. However, this is a too negative point of view and one that does not take into account that, of course, the integrations cannot be carried out in practice. The value of Eq. (6) lies therein that it provides us with an *exact* formal definition of the effective interaction, on which approximations of varying and, hopefully, increasing accuracy can be applied. Moreover, an effective interaction sheds light on the effects that the coarse-grained degrees of freedom have on the physics of the ones that remain in the description and, in this way, it also makes direct contact with experimental observations.[1]

In the sections to follow, we will proceed with a more detailed analysis of various aspects of the effective potential, of particular properties of its structure depending on the physical systems at hand and of suitable approximation methods. Before doing so, let us set forth a few important points regarding important limitations it has and relevant difficulties associated with this theoretical tool.

1. The process of the coarse-graining generates, first of all, one additional, extensive term in the effective Hamiltonian, denoted $\mathcal{H}_0$ in Eq. (10). Although this has no effect on the structural functions of γ-type-particles, it does influence the free energy, in which it appears as an additive constant that depends, however, on the mixture composition and on the nature of its constituent units.

2. Information on the densities of the species that have been coarse-grained, on the temperature and on the interactions they have among one another and with the remaining γ-species are all encoded in the effective potential $U_{\rm eff}$. This makes the latter *tunable* by suitable choice of the composition and physical nature of the mixture.

3. The effective potential $U_{\rm eff}$ is by no means pairwise additive. This is the price we have to pay for simplifying the system. This feature is strongly reminiscent of the problem of *proliferation of interactions* encountered when performing a (real-space), block-spin type renormalization-group coarse-graining of lattice Hamiltonians;[9] note that also in that case we obtain extensive, additive terms akin to $\mathcal{H}_0$ in the renormalized Hamiltonian. Still, a pairwise-additivity approximation can often be applied and it turns out to be quite accurate in some occasions. In the sections to follow, we will

see examples of pairwise-additivity approximations that range in their quality from excellent to hopeless.

4. The coarse-graining process does result into loss of information. In particular, any correlation function or any quantity that involves the coordinates of the coarse-grained degrees of freedom cannot be calculated any more within the framework of $\mathcal{H}_{\rm eff}$.

5. The general approach to coarse-graining was presented above in the canonical ensemble. Depending on the physical situation at hand, working in different ensembles, such as the semi-grand or the grand canonical ones, when possible, could be advantageous.

Having thus set the stage for the effective Hamiltonian, we proceed in what follows with a more and more detailed examination of particular cases and physical systems.

2.2 Unconstrained systems

We consider for simplicity a *binary* mixture of two species of particles, 1 and 2, with populations N_1, N_2, and coordinates $\{\mathbf{R}_i\}$, $i = 1, 2, \ldots, N_1$, and $\{\mathbf{r}_j\}$, $j = 1, 2, \ldots, N_2$, respectively. We are further assuming that the constituent particles of the two species have no internal degrees of freedom and that the can be represented as point particles. In the opposite case in which, e.g., each species is some complex macromolecule, the considerations of this section should be preceded by the coarse-graining of the internal degrees of freedom and the substitution of the whole macromolecular aggregate by a single effective coordinate, as described in Section 2.3 that follows. Setting the masses of the two species to be m_1 and m_2, the Hamiltonian $\mathcal{H}$ of the mixture is expressed as the sum of the kinetic and interaction parts, $\mathcal{H} = K + U$. The kinetic energy is trivially expressed as:

$$K = \sum_{i=1}^{N_1} \frac{\mathbf{P}_i^2}{2m_1} + \sum_{j=1}^{N_1} \frac{\mathbf{p}_j^2}{2m_2}, \tag{11}$$

with the canonical momenta $\mathbf{P}_i$ and $\mathbf{p}_j$ of the two species. The interaction part, on the other hand, is conveniently separated into a sum of intra- and interspecies interactions, $U_{\rm intra} = U_{11} + U_{22}$ and $U_{\rm inter} = U_{12}$ respectively, whereby:

$$U_{11} = \sum_{i=1}^{N_1} \sum_{j>i}^{N_1} \phi_{11}(|\mathbf{R}_i - \mathbf{R}_j|), \tag{12}$$

$$U_{22} = \sum_{i=1}^{N_2} \sum_{j>i}^{N_2} \phi_{22}(|\mathbf{r}_i - \mathbf{r}_j|), \tag{13}$$

and

$$U_{12} = \sum_{i=1}^{N_1} \sum_{j=i}^{N_2} \phi_{12}(|\mathbf{R}_i - \mathbf{r}_j|), \tag{14}$$

introducing the microscopic pair potentials $\phi_{ij}(x)$, which are assumed to be isotropic.[a]

[a]This is not a restriction: the formalism can also be extended to anisotropic interactions but we refrain from examining this case in the present manuscript.

A comprehensive theoretical approach to coarse-graining such a system in eliminating component 2 and obtaining as a result an effective Hamiltonian for component 1 only was set forward by Dijkstra *et al.*;[10] we summarize below the main results of this approach and we discuss their physical meaning. To begin with, the most suitable statistical ensemble to treat this problem turns out to be not the canonical one, in which the particle numbers N_1, N_2 are fixed together with temperature T and volume V but rather the semigrand one, in which the particle number N_2 is allowed to fluctuate via a fixed chemical potential μ_2 of the second species. The corresponding thermodynamic potential in this ensemble is the semigrand free energy $J(N_1, \mu_2, V, T)$, which is expressed as the Legendre transformation of the Helmholtz free energy $F(N_1, N_2, V, T)$, viz:

$$J(N_1, \mu_2, V, T) = F(N_1, N_2, V, T) - \mu_2 N_2. \tag{15}$$

The semigrand free energy is, of course, $-k_{\mathrm{B}}T$ times the logarithm of the semigrand partition function, for which a sum upon all particle numbers N_2 is performed:

$$\exp(-\beta J) = \sum_{N_2=0}^{\infty} \exp\left[-\beta\left(F - \mu_2 N_2\right)\right], \tag{16}$$

whereas

$$\exp(-\beta F) = \frac{1}{N_1! {\Lambda_1}^{3N_1}} \frac{1}{N_2! {\Lambda_2}^{3N_2}} \mathrm{Tr}_1 \mathrm{Tr}_2 \exp(-\beta H). \tag{17}$$

From Eqs. (15) and (16), and setting the activity of species 2 as $z_2 = \Lambda_2^{-3}\exp(\beta\mu_2)$, we immediately obtain an effective, one-component version of the semigrand partition function as

$$\exp(-\beta J) = \frac{1}{N_1! \Lambda_1^{3N_1}} \mathrm{Tr}_1 \exp\left[-\beta\left(U_{11} + \Omega\right)\right], \tag{18}$$

where Ω is the constrained grand partition function of the particles of species 2 under the condition that those of species 1 are held fixed at positions $(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_1})$:

$$\Omega(\{\mathbf{R}_i\}; N_1, z_2, V, T) = \sum_{N_2=0}^{\infty} \frac{z_2^{N_2}}{N_2!} \mathrm{Tr}_2 \exp\left[-\beta\left(U_{12} + U_{22}\right)\right]. \tag{19}$$

Following the notation introduced in Section 2.1, the effective potential function $\tilde{U}_1$ is expressed as the sum $U_{11} + \Omega$; the former quantity is the direct interactions between particles of species 1, whereas the latter represents the interactions that are mediated through the second component:

$$\tilde{U}_1(\{\mathbf{R}_i\}; N_1, z_2, V, T) = U_{11}(\{\mathbf{R}_i\}) + \Omega(\{\mathbf{R}_i\}; N_1, z_2, V, T). \tag{20}$$

Of particular importance for the case of unconstrained systems is the remarkable fact that the induced interaction Ω can be systematically decomposed into a sum of n-body terms, $n = 0, 1, 2, \ldots$, by means of diagrammatic expansions based on 1-2 and 2-2 Mayer-function bonds defined as

$$\begin{aligned} f(|\mathbf{R}_i - \mathbf{r}_j|) &= \exp\left[-\beta\phi_{12}(|\mathbf{R}_i - \mathbf{r}_j|)\right] - 1, \\ g(|\mathbf{r}_i - \mathbf{r}_j|) &= \exp\left[-\beta\phi_{22}(|\mathbf{r}_i - \mathbf{r}_j|)\right] - 1. \end{aligned} \tag{21}$$

For the details of the diagrammatic expansion, we refer the reader to the original publication.[10] To illustrate the power of the formalism and the main results to which it leads, let us further define a subset of the cross-interaction term U_{12}, call it $U_{12}^{(n)}$, which involves *precisely n* particles of species 1 interacting with N_2 particles of species 2, namely

$$U_{12}^{(n)}(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_n) = \sum_{i=1}^{n} \sum_{j=i}^{N_2} \phi_{12}(|\mathbf{R}_i - \mathbf{r}_j|), \tag{22}$$

where, evidently, $U_{12} \equiv U_{12}^{(N_1)}$, cf. Eq. (14), and we have dropped from the list of arguments the coordinates $\{\mathbf{r}_j\}$ for parsimony in the notation (in what follows these will be integrated upon anyway). The induced interaction $\Omega(\{\mathbf{R}_i\}; N_1, z_2, V, T)$ can then be decomposed as a sum of n-body terms, each one involving thermodynamic averages that involve, respectively, *only* the n-body cross-interaction potentials $U_{12}^{(n)}$ defined above:

$$\Omega(\{\mathbf{R}_i\}; N_1, z_2, V, T) = \sum_{n=0}^{N_1} \Omega_n(\{\mathbf{R}_i\}; N_1, z_2, V, T). \tag{23}$$

It is particularly instructive to look into some more detail into the first three terms, $n = 0, 1$ and $n = 2$. The zeroth-order term, Ω_0 involves a system that contains no particles of species 1 at all, and it is nothing else than the grand potential of a pure system of particles of type 2 at fixed chemical potential μ_2. Accordingly, it is given as

$$\Omega_0(z_2, V, T) = -Vp(z_2, T), \tag{24}$$

where $p(z_2, T)$ is the pressure of a pure system of particles of species 2 with activity z_2. Since we are working at the semigrand ensemble, this term is, at the same time, the pressure of a reservoir of particles 2, which is at particle exchange equilibrium with the mixture at hand through the presence of a semipermeable membrane that allows the passage of particles of species 2 only, blocking the particles of species 1. The Ω_1-term formally involves a single particle of species 1 inserted in the bath of particles of species 2. Through the homogeneity (translational invariance) of the bath of species 2, the dependence on the point of insertion disappears and Ω_1 also becomes an extensive contribution which does not depend on the coordinates of species 1. In particular,

$$\Omega_1(N_1, z_2, T) = N_1 \omega_1(z_2, T), \tag{25}$$

where

$$\omega_1(z_2, T) = -k_{\mathrm{B}} T \ln \left\langle \exp\left[-\beta U_{12}^{(1)}(\mathbf{0})\right] \right\rangle_{z_2}, \tag{26}$$

where the notation $\langle \cdots \rangle_{z_2}$ denotes a grand canconical expectation value taken at the reservoir-Hamiltonian, i.e., within a system interacting by means of U_{22} only. In other words, in Eq. (26), the particle of species 1 is inserted, without loss of generality, at the origin $\mathbf{0}$ and it acts as a ghost particle: it neither displaces particles of the bath nor does it affect their configuration. We thus have to perform a measurement of the quantity in the brackets without the quantity itself being active in determining physical correlation the system, which are dictated by U_{22} *alone*. One immediately recognizes here the famous Widom insertion[11] and Eq. (26) expresses $\omega_1(z_2, T)$ as one contribution of the sea of species 2 to the excess chemical potential of the particles of species 1.

Taken together, the terms Ω_0 and Ω_1 thus constitute the configuration-independent, extensive contribution $\mathcal{H}_0$ to the effective Hamiltonian, cf. Eq. (10):

$$\mathcal{H}_0 = -Vp_2(z_2, T) + N_1\omega_1(z_2, T), \tag{27}$$

the interaction terms involving all the higher-order contributions to Ω:

$$U_{\text{eff}}(\{\mathbf{R}_i\}; N_1, z_2, T) = U_{11}(\{\mathbf{R}_i\}) + \sum_{n=2}^{N_1} \Omega_n(\{\mathbf{R}_i\}; N_1, z_2, T). \tag{28}$$

Eq. (28) manifests what has been mentioned previously: the effective interaction inadvertently generates n-body coupling terms in the Hamiltonian, going thus beyond pair-additivity. Still, effective pair potentials usually dominate and it is useful to look in more detail into the structure of the term $\Omega_2(\{\mathbf{R}_i\}; N_1, z_2, T)$. In analogy to the bare interaction $U_{11}(\{\mathbf{R}_i\})$, Eq. (12), this term is pair-decomposable as

$$\Omega_2(\{\mathbf{R}_i\}; N_1, z_2, T) = \sum_{i=1}^{N_1} \sum_{j>i}^{N_1} \omega_2(|\mathbf{R}_i - \mathbf{R}_j|; z_2, T). \tag{29}$$

Setting $R \equiv |\mathbf{R}_i - \mathbf{R}_j|$, the species-2-induced pair interaction potential $\omega_2(R; z_2, T)$ is expressed as

$$\omega_2(R; z_2, T) = -k_{\text{B}}T \ln \left[\frac{\left\langle \exp\left[-\beta U_{12}^{(2)}(\mathbf{R}_i, \mathbf{R}_j)\right]\right\rangle_{z_2}}{\left\langle \exp\left[-\beta U_{12}^{(1)}(\mathbf{0})\right]\right\rangle_{z_2}^2} \right]. \tag{30}$$

The reduction of the dependence from $\mathbf{R}_i$, $\mathbf{R}_i$ to just the magnitude of their difference is the result of the averaging in an isotropic system.

It is worth taking a closer look at the denominator of Eq. (30) above. It expresses the expectation value of the Boltzmann-weighted cost of introducing a particle of type 1 in a sea of particles 2, *squared*. In other words, it can also be looked upon as the same cost for the introduction of *two* particles of type 1, separated by an infinite distance from one another. Certainly, if the two are infinitely far apart, the insertion of one of them is uncorrelated from the insertion of the other. Assuming, therefore, that one of the particles is inserted at the origin $\mathbf{0}$ and the other at $\mathbf{S}$, we can write

$$\begin{aligned}
\left\langle \exp\left[-\beta U_{12}^{(1)}(\mathbf{0})\right]\right\rangle_{z_2}^2 &= \left\langle \exp\left[-\beta U_{12}^{(1)}(\mathbf{0})\right]\right\rangle_{z_2} \left\langle \exp\left[-\beta U_{12}^{(1)}(\mathbf{S})\right]\right\rangle_{z_2} \\
&= \left\langle \exp\left[-\beta U_{12}^{(1)}(\mathbf{0})\right] \exp\left[-\beta U_{12}^{(1)}(\mathbf{S})\right]\right\rangle_{z_2, \mathbf{S}\to\infty} \\
&= \left\langle \exp\left[-\beta \left(U_{12}^{(1)}(\mathbf{0}) + U_{12}^{(1)}(\mathbf{S})\right)\right]\right\rangle_{z_2, \mathbf{S}\to\infty} \\
&= \left\langle \exp\left[-\beta U_{12}^{(2)}(\mathbf{0}, \mathbf{S})\right]\right\rangle_{z_2, \mathbf{S}\to\infty}.
\end{aligned} \tag{31}$$

In the first line above, we used the independence of the expectation value on the point of insertion. In the second one, we took the limit $\mathbf{S} \to \infty$ and made use of the fact that in this case the two random insertions yield costs uncorrelated to one another, and thus

the expectation value of the product coincides with the product of the expectation values. Combining Eqs. (30) and (31), we can now express the induced pair potential as

$$\begin{aligned}\omega_2(R; z_2, T) = &- k_{\mathrm{B}}T \ln \left\langle \exp\left[-\beta U_{12}^{(2)}(\mathbf{0}, \mathbf{R})\right] \right\rangle_{z_2} \\ &+ k_{\mathrm{B}}T \ln \left\langle \exp\left[-\beta U_{12}^{(2)}(\mathbf{0}, \mathbf{R} \to \infty)\right] \right\rangle_{z_2} . \end{aligned} \tag{32}$$

The induced interaction $\omega_2(R; z_2, T)$ is therefore a difference between two semigrand free energy costs: the one that obtains by the (virtual) insertion of two type-1 particles, separated by $\mathbf{R}$, in a sea of small ones *minus* the same quantity when these particles are inserted at places infinitely far apart from one another. The procedure a generalized Widom insertion, in which two particles are virtually placed inside the system instead of just one. The induced interaction evidently fulfills the property $\omega_2(R \to \infty; z_2, T) \to 0$.

Eq. (32) above is exact. Its value lies therein that it gives us a point of reference, a theoretical milestone to which we can refer when applying any further approximations to calculate the induced interaction, a quantity which, in the realm of the coarse-graining of binary, unconstrained systems is also called *depletion interaction*.[12–15, 10, 16] However, its practical value for dense systems is very limited. To illustrate this point, imagine of a concrete physical system, i.e., a mixture of hard spheres, large (species 1) and small (species 2). Suppose that the small system is very dense and keep in mind that the expectation values of Eq. (32) are taken in an ensemble of pure small hard sphere systems, which are unaware of the existence of the large ones; the insertions are, as mentioned above, *virtual*. What is the chance that such a system will spontaneously create a void sufficiently large to accommodate one and two large hard spheres? It is ridiculously small! In the vast majority of cases, the attempt to insert one or two large hard spheres will result in overlaps with the small ones and both terms in the exponents of Eq. (32) will have the value *minus infinity*. What Eq. (32) actually does, is to take the difference between the logarithms of two numbers that are both extremely small; the difference, however, is a reasonable number of order unity.[13] Attempting to do this in the way Eq. (32) suggests is hopeless at high densities of the small component, if the steric interactions with the big one are strong. Accordingly, one has to resort to other methods, in which the large particles are inserted in the fully interacting system *from the outset*. If this is the case, then the relative difference of the two logarithms (or, equivalently, the ratio of the two expectation values) is measured directly, and therefore the problem is circumnavigated in this way. We will return to this point in Section 3 below.

Finally, we briefly mention the higher-order interactions, taking the three-body term $\Omega_3(\{\mathbf{R}\}; N_1, z_2, T)$ as a case in point: This term can be decomposed as

$$\Omega_3(\{\mathbf{R}\}; N_1, z_2, T) = \sum_{i=1}^{N_1} \sum_{j>i}^{N_1} \sum_{k>j}^{N_1} \omega_3(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k; z_2, T), \tag{33}$$

with the three-body induced potential taking the form:

$$\omega_3(\mathbf{R}_i,\mathbf{R}_j,\mathbf{R}_k;z_2,T) = -\,k_{\mathrm{B}}T\ln\frac{\left\langle\exp\left[-\beta U_{12}^{(3)}(\mathbf{R}_i,\mathbf{R}_j,\mathbf{R}_k)\right]\right\rangle_{z_2}}{\left\langle\exp\left[-\beta U_{12}^{(1)}(\mathbf{0})\right]\right\rangle_{z_2}^{3}} + k_{\mathrm{B}}T\sum_{(m,n)}\ln\frac{\left\langle\exp\left[-\beta U_{12}^{(2)}(\mathbf{R}_m,\mathbf{R}_n)\right]\right\rangle_{z_2}}{\left\langle\exp\left[-\beta U_{12}^{(1)}(\mathbf{0})\right]\right\rangle_{z_2}^{2}}, \tag{34}$$

where the sum at the second line runs over the combinations $(m,n)=(i,j),(j,k),(k,i)$. The structure of Eq. (34) reveals the physical meaning of the three-body induced potential $\omega_3(\mathbf{R}_i,\mathbf{R}_j,\mathbf{R}_k;z_2,T)$. In the first term, three particles of species 1 are Widom-inserted and the cost of this insertion *relative* to the case in which all three are infinitely far away from one another is calculated. The second term is, evidently, *minus* the sum of the three two-body interactions between the particles, cf. Eq. (30). Thus, $\omega_3(\mathbf{R}_i,\mathbf{R}_j,\mathbf{R}_k;z_2,T)$ measures *precisely* the *deviations* of the free energy cost of insertion of three particles from those that would result from pair additivity. Such deviations are expected on physical grounds, if we think of three large hard spheres being *really* and not just virtually inserted in a sea of small ones. The distortions they cause on the density profile of the small spheres are not necessarily the superposition of the distortions caused by insertions in pairs. There are genuine three-body effects, which are not reducible to superpositions of pairs. This is a general characteristic of the effective, as opposed to fundamental (i.e., microscopic) interactions, such as the Coulomb potential. Even higher-order induced potentials follow the same logic; thus, Ω_n counts the genuine n-body free energy insertion cost of n particles, which is *not* included in all lower-order terms. Their functional form follows from this statement and we will therefore not delve into this topic in any further detail.

Gathering the above results the effective, single-species Hamiltonian takes the form:

$$\mathcal{H}_{\mathrm{eff}}(\{\mathbf{P}_i\},\{\mathbf{R}_i\};V,N_1,z_2,T) = -\,Vp(z_2,T)+N_1\omega_1(z_2,T) + \sum_{i=1}^{N_1}\frac{\mathbf{P}_i^2}{2m_1}+\sum_{i=1}^{N_1}\sum_{j>i}^{N_1}\left[\phi_{11}(R_{ij})+\omega_2(R_{ij};z_2,T)\right] + \sum_{i=1}^{N_1}\sum_{j>i}^{N_1}\sum_{k>j}^{N_1}\omega_3(\mathbf{R}_i,\mathbf{R}_j,\mathbf{R}_k;z_2,T)+\cdots \tag{35}$$

Excluding the two first terms on the right-hand side from the statistical mechanics of the system, we would obtain an apparent pressure $\Pi(\rho_1,z_2,T)$ for the system as well as an apparent chemical potential $\mu_1'(\rho_1,z_2,T)$. From Eq. (35), it follows that the total pressure $P(\rho_1,z_2,T)$ of the system and the total chemical potential $\mu_1(\rho_1,z_2,T)$ of species 1 are given by the expressions:

$$P(\rho_1,z_2,T)=\Pi(\rho_1,z_2,T)+p(z_2,T); \tag{36}$$

$$\mu_1(\rho_1,z_2,T)=\mu_1'(\rho_1,z_2,T)+\omega_1(z_2,T). \tag{37}$$

Accordingly, the quantity $\Pi(\rho_1,z_2,T)$ is recognized as the osmotic pressure across a semipermeable membrane separating the system from a reservoir of species 2, which allows the exchange of this species only, setting its activity to z_2 for both the system and the

reservoir. The reader should confirm that the terms $p(z_2, T)$ and $\omega_1(z_2, T)$ do not affect the location of possible phase boundaries $(\rho_1^{\mathrm{A}}, \rho_2^{\mathrm{B}})$ for the densities of species 1 between two coexisting phases A and B.

We conclude this Section with a remark regarding the correlations between particles of species 1, which are immersed in a fluid of species 2. As discussed above, the latter is described by its partial chemical potential μ_2, or equivalently by its activity z_2, which can be thought of as that of a reservoir of the second species at the same chemical potential. Let us truncate the series of Eq. (35) at the pair potential level, for simplicity. Then, the total pair interaction potential, $\phi_{\mathrm{eff}}(r; z_2, T)$ is given as

$$\phi_{\mathrm{eff}}(r; z_2, T) = \phi_{11}(r) + \omega_2(r; z_2, T). \tag{38}$$

The radial distribution function[17] $g_{11}(r; \rho_1, z_2, T)$ between particles of species 1 depends parametrically on the number density $\rho_1 = N_1/V$ of the same, as well as on z_2 and the temperature; the z_2-dependence comes from the induced interaction $\omega_2(r; z_2, T)$. All techniques from the theory of the liquid state to calculate the correlation functions can be applied to the system interacting by means of the effective potential $\phi_{\mathrm{eff}}(r)$. In particular, at the limit of vanishing density of species 1, it holds:

$$\lim_{\rho_1 \to 0} g_{11}(r; \rho_1, z_2, T) = \exp\left[-\beta \phi_{\mathrm{eff}}(r; z_2, T)\right]. \tag{39}$$

Eq. (39) can be very useful for measuring the effective potential. We will return to this point in Section 3.

2.3 Constrained systems

Another broad class of soft matter systems for which coarse-graining is involved have mutual *constraints* in their number fluctuations, and therefore the semigrand ensemble formalism set forth in the preceding section cannot be applied in this case. Two categories of such common systems exist, namely the following.

1. Mixtures of entities that carry electric charge, such as ionic mixtures, liquid metals and charged colloidal suspensions, containing in general salt.[18] In this case, one is dealing with a ν-component mixture, each of the constituent particles carrying charge $Z_\nu e$. Since electroneutrality must be strictly obeyed so that the thermodynamic limit exists, the number densities ρ_ν of the components are bound to fulfill the condition:

$$\sum_{i=1}^{\nu} \rho_i Z_i = 0. \tag{40}$$

In most cases, we are dealing with two-component mixtures, so that Eq. (40) above immediately implies that the particle number N_2 cannot fluctuate independently of N_1; the two are mutually constrained.

2. Systems that consist of macromolecular aggregates with internal fluctuations, each aggregate $\mathcal{I}$ consisting of a fixed number of $M_{\mathcal{I}}$ microscopic units; a typical example are polymers of different architectures (chain, ring, star, brush, dendritic, cross-linked etc.) Here, one wishes to represent the full object in a coarse-grained fashion by reducing the $M_{\mathcal{I}}$ degrees of freedom to a much smaller number, typically just one

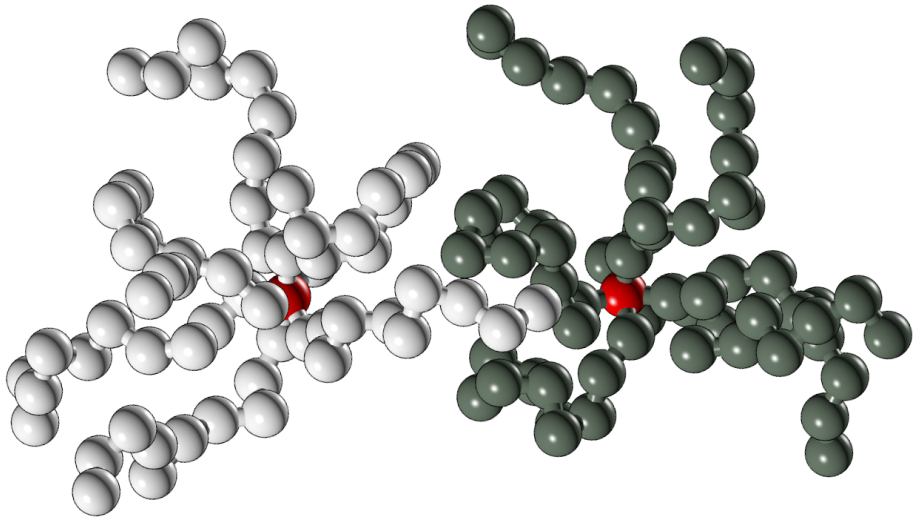


Figure 1: Simulation snapshots of two model star polymers at moderate mutual overlap. The spheres denote monomers along the arms, whereas the two red balls are the centers of the stars, which are chosen as effective coordinates, as explained in the text.

effective coordinate. Evidently, the number of effective coordinates is bound to the number of the microscopic ones by the architecture (chemistry), which is a *quenched* property of the system. The microscopic and the coarse-grained numbers of degrees of freedom cannot fluctuate independently of one another.

The first class of systems constitute one of the most well-studied category of colloidal suspensions and one that still is very much an active topic of research. Among the most celebrated results of the coarse-graining approach to charge-stabilized colloidal suspensions are the Poisson-Boltzmann theory and its linearized version, known as Debye-Hückel theory, as well as the associated Derjaguin-Landau-Verwey-Overbeek (DLVO) potential. We will not discuss any further these system at the present Chapter, referring the reader to some of the most representative articles on this subject in the literature.[19] We focus, instead on the second class of systems. The formalism for deriving an effective interaction is set in the framework of the canonical ensemble, as shown in Section 2.1. In what follows, we will make the derivation tractable by considering *just two* macromolecules and formally tracing out the microscopic degrees of freedom, so that a pair effective interaction between the mesoscopic degrees of freedom remains.

We consider two star polymers with f arms, see Figure 1, as a representative example. The average spherical symmetry of the monomer density profile around the central, anchoring point of the arms, makes the choice of the latter as an effective coordinate natural; however, the discussion is general and any other material coordinate or even the center of mass can be chosen as such. The goal of the coarse-graining is thus to canonically average out all the degrees of freedom of the monomers along the arms and be left with an effective interaction $\phi_{\rm eff}(R)$ between the centers, whereby the first star is placed at the origin and the second at a position $\mathbf{R}$ from it.

We assume that each star contains $M+1$ beads, one being the central one and M along the arms.[b] We denote the position of the i-th bead of the α-star as $\mathbf{s}_{i\alpha}$, $i = 0, 1, \ldots, M$ being the monomer index and $\alpha = 1, 2$ the molecule index, whereby $\mathbf{s}_{01}$ and $\mathbf{s}_{02}$ are the coordinates of the two central particles. The total potential energy U of the two interacting stars can be written as the sum of intramolecular, $U_{\alpha\alpha}$, and intermolecular, $U_{\alpha\beta}$, $\alpha \neq \beta$ and $\alpha, \beta = 1, 2$, contributions as

$$\begin{aligned} U(\{\mathbf{s}_{i1}\}, \{\mathbf{s}_{j2}\}) &= U_{11}(\{\mathbf{s}_{i1}\}) + U_{22}(\{\mathbf{s}_{j2}\}) + U_{12}(\{\mathbf{s}_{i1}\}, \{\mathbf{s}_{j2}\}) \\ &\equiv U_{\text{intra}}(\{\mathbf{s}_{i1}\}, \{\mathbf{s}_{j2}\}) + U_{12}(\{\mathbf{s}_{i1}\}, \{\mathbf{s}_{j2}\}), \end{aligned} \tag{41}$$

where we have grouped together the two intramolecular terms in U_{intra}, which is free of interactions between the stars. The intramolecular potential energy functions contain contributions from both bonded and non-bonded monomers, whereas, evidently, the inrermolecular one features exclusively non-bonded terms. Let us denote as $\mathcal{W}_2$ the configuration part of the partition function of the two stars, expressed as [cf. Eq. (3)]:

$$\mathcal{W}_2 = \int \mathrm{d}\mathbf{s}_{01} \cdots \int \mathrm{d}\mathbf{s}_{M2} \exp\left[-\beta U(\{\mathbf{s}_{i1}\}, \{\mathbf{s}_{j2}\})\right]. \tag{42}$$

Moreover, let $\mathcal{W}_1$ be the partition function of a single star with its center fixed at some position in space, namely

$$\mathcal{W}_1 = \int \mathrm{d}\mathbf{s}_{1\alpha} \cdots \int \mathrm{d}\mathbf{s}_{M\alpha} \exp\left[-\beta U_{\alpha\alpha}(\{\mathbf{s}_{i\alpha}\})\right], \tag{43}$$

where, in the last equation, it is irrelevant whether $\alpha = 1$ or $\alpha = 2$, since the two molecules are identical.

The key quantity is a *constrained Helmholtz free energy* $F(\mathbf{R})$ defined as the partial trace over all other degrees of freedom, with the two centers kept at positions $\mathbf{s}_{01} = \mathbf{0}$ and $\mathbf{s}_{02} = \mathbf{R}$, viz.:

$$\exp\left[-\beta F(\mathbf{R})\right] = \int \mathrm{d}\mathbf{s}_{01} \cdots \int \mathrm{d}\mathbf{s}_{M2} \exp\left[-\beta U(\{\mathbf{s}_{i1}\}, \{\mathbf{s}_{j2}\})\right] \delta\left(\mathbf{s}_{01}\right) \delta\left(\mathbf{s}_{02} - \mathbf{R}\right). \tag{44}$$

With $p(\mathbf{R})$ denoting the probability density of finding the center of the second star at position $\mathbf{R}$ relative to the first, it holds by construction of the quantity $F(\mathbf{R})$ that:

$$p(\mathbf{R}) \propto \exp\left[-\beta F(\mathbf{R})\right], \tag{45}$$

the constant of proportionality being the inverse of the partition function, $\mathcal{W}_2^{-1}$, which is irrelevant as long as we are interested in the *relative* probabilities:

$$\frac{p(\mathbf{R})}{p(\mathbf{R}_0)} = \exp\left[-\beta\left(F(\mathbf{R}) - F(\mathbf{R}_0)\right)\right], \tag{46}$$

where $\mathbf{R}_0$ is some arbitrary reference vector. Of particular usefulness is the choice $\mathbf{R}_0 \to \infty$. In this case, the quantity $\exp\left[-\beta F(\mathbf{R}_0)\right]$, Eq. (44), takes a simple form. Indeed, since every molecule has a finite spatial extension (of the order of its gyration radius), when the centers of the two are taken infinitely far apart, all intermolecular contributions

[b] Nothing changes in the formalism to follow if the two stars have different monomer numbers but we set them equal for simplicity in notation.

U_{12} vanish and the expression in Eq. (44) reduces to the product of the partition functions of two isolated stars, each with its center fixed in space, Eq. (43):

$$\exp\left[-\beta F(\mathbf{R}_0 \to \infty)\right] = \mathcal{W}_1^2. \tag{47}$$

We now *define* the effective interaction $\phi_{\text{eff}}(\mathbf{r})$ between the two macromolecules as

$$\phi_{\text{eff}}(\mathbf{r}) = F(\mathbf{r}) - F(\mathbf{R}_0 \to \infty). \tag{48}$$

Combining Eqs. (46) and (48), we readily obtain

$$\exp\left[-\beta\phi_{\text{eff}}(\mathbf{r})\right] = \frac{p(\mathbf{r})}{p(\mathbf{R}_0 \to \infty)}. \tag{49}$$

However, according to the discussion preceding Eq. (47), as $\mathbf{R}_0 \to \infty$, the two stars become uncorrelated to one another. Accordingly, $p(\mathbf{R}_0 \to \infty) = p_0(\mathbf{r})$ for any $\mathbf{r}$, where $p_0(\mathbf{r})$ is the probability density in the Hamiltonian U_0 of two noninteracting stars; the latter form an ideal gas and therefore

$$\frac{p(\mathbf{r})}{p(\mathbf{R}_0 \to \infty)} = \frac{p(\mathbf{r})}{p_0(\mathbf{r})} \equiv g(r), \tag{50}$$

introducing the infinite-dilution radial distribution function $g(r)$ of the centers of *only two* star polymers in the volume V, in the absence of any others. Accordingly, from Eqs. (49) and (50) we obtain that the effective interaction is isotropic and it can be expressed as

$$\lim_{\rho\to 0} g(r;\rho) = \exp\left[-\beta\phi_{\text{eff}}(r)\right], \tag{51}$$

which is identical to Eq. (39).

Contrary to the discussion in Section 2.2, we have not presented here a general scheme that generates at once expressions for all n-body interactions. However, the generalization to $n > 2$ is conceptually straightforward: one considers the constrained free energy $F(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_n)$ of the centers of n macromolecules and subtracts from it the sum of all contributions of the immediately lower, $(n-1)$-order effective potential, defining in this way the n-body interaction.[20] Depending on the physical system, and in particular on the deformability of the macromolecules at hand,[21,6,22] the higher-order interactions can become very important. We will return to this point in Section 6.

3 Methods of Calculating Effective Pair Potentials

3.1 Methods based on the radial distribution function

The most straightforward way for calculating the effective interaction in a simulation is offered by Eq. (39) or Eq. (51). Ideally, one would perform a standard Monte Carlo (MC) or Molecular Dynamics (MD) simulation of just two macromolecular aggregates in a sea of smaller particles [Eq. (39)] or just by themselves [Eq. (51)] and measure, by calculating the properly normalized histogram of separations r the infinite-dilution limit radial distribution function $g(r;\rho \to 0)$; the limit $\rho \to 0$ is, of course, an idealization but having just two particles in a sufficiently large simulation box is sufficient. From that, the effective interaction can be obtained as

$$\phi_{\text{eff}}(r) = -k_{\text{B}}T \ln g(r;\rho \to 0). \tag{52}$$

The simplicity of this approach is deceptive. Regions in which the effective interaction attains values significantly higher than $k_{\rm B}T$ will be rarely visited in an unbiased simulation and therefore sampling will be extremely poor there. In a similar spirit, if $\phi_{\rm eff}(r)$ has deep negative minima, separated from the rest by high barriers, in a free simulation the two particles will be trapped or 'locked' into these and, again, it will be very hard to explore efficiently and accurately the whole range of separations r. One possibility to overcome this problem is to add a bias potential,[23] $\Phi_{\rm bias}(r)$, to the interaction and measure the modified radial distribution function, $\tilde{g}(r, \rho \to 0)$ given by

$$\phi_{\rm eff}(r) + \Phi_{\rm bias}(r) = -k_{\rm B}T \ln \tilde{g}(r; \rho \to 0). \tag{53}$$

Ideally, one would like to choose $\Phi_{\rm bias}(r)$ in such a way that all separations r are sampled uniformly, i.e., $\tilde{g}(r; \rho \to 0) = 1$, which immediately implies, through Eq. (53), $\Phi_{\rm bias}(r) = -\phi_{\rm eff}(r)$. One would have to add a bias potential opposite to the unknown effective interaction, which, if we knew in advance, we wouldn't need to make a simulation to find out what it is! Evidently, we have to resort to some more sophisticated approaches, in which the bias potential is determined in an iterative way[24] or it is built-up locally during the simulation in a Wang-Landau-type simulation scheme.[25]

To avoid iterative approaches, one may choose the bias potential to be *local*, constraining thereby r within smaller regions or 'windows'. The technique is thereby called 'window sampling' and the basic idea is the following: the interval $r \in [r_{\rm min}, r_{\rm max}]$ is separated into sub-intervals or windows of width w, which may vary from one window to the other. Within each window, the effective potential $\phi_{\rm eff}(r)$ is supposed or expected to vary by at most a few $k_{\rm B}T$, so that when r is restricted to lie within any given window, statistics there is good and the whole width w is sampled well. Common choices for the local bias potential are an infinite well that vanishes within w and diverges elsewhere[26,27] or a harmonic potential $\Phi_{\rm bias}(r) = k_j(r - r_j)^2$ centered at the midpoint r_j of the window. Separate histograms $\tilde{g}^{(j)}(r)$ within each of these windows are collected and at the end of the simulation the effective interaction results by merging together the pieces $\phi_{\rm eff}^{(j)}(r)$ obtained in each window:

$$\phi_{\rm eff}^{(j)}(r) = -k_{\rm B}T \ln \tilde{g}^{(j)}(r) - \Phi_{\rm bias}(r) + c_j, \tag{54}$$

where the c_j's are arbitrary constants that are used to match the pieces of the effective potential. In particular, the various pieces are 'vertically shifted' via appropriate choices of the constants c_j so as to obtain a smooth function $\phi_{\rm eff}(r)$. To achieve this, some overlap between the windows is necessary.[26] Alternatively, one can resort to more sophisticated methods, such as histogram reweighting[28,29] to obtain a smooth 'patching' of the $\phi_{\rm eff}^{(j)}(r)$ into $\phi_{\rm eff}(r)$.

3.2 Methods based on the effective force

Instead of measuring directly the effective potential $\phi_{\rm eff}(r; z_2, T)$, one can resort to measuring its gradient, which is the effective force $\mathbf{F}_{\rm eff}(r)$, namely

$$\mathbf{F}_{\rm eff}(r; z_2, T) = -\nabla_r \phi_{\rm eff}(r; z_2, T). \tag{55}$$

In what follows, we will derive expressions for the effective force for both unconstrained [Section 2.2] and constrained [Section 2.3] systems.

3.2.1 Unconstrained systems

Our starting point is Eq. (32). Consider the first term in the brackets at the right-hand side of Eq. (32). Using Eq. (22) and setting $\mathbf{R} \to \mathbf{r}$, we have:

$$\left\langle \exp\left[-\beta U_{12}^{(2)}(\mathbf{0},\mathbf{r})\right]\right\rangle_{z_2} = \frac{1}{\Xi_2(z_2,T,V)}\sum_{N_2=0}^{\infty}\frac{z_2^{N_2}}{N_2!} \times \mathrm{Tr}_2\,\exp\left\{-\beta\left(U_{22}(\{\mathbf{r}_j\}) + \sum_{j=1}^{N_2}\left[\phi_{12}(r_j) + \phi_{12}(|\mathbf{r}-\mathbf{r}_j|)\right]\right)\right\}, \tag{56}$$

where $\Xi_2(z_2,T,V)$ is the grand partition function of the reservoir of particles 2. The term in the parentheses in the exponential of the right-hand-side of Eq. (56) above can be interpreted as the configurational part of a modified Hamiltonian, in which two type-1 particles, held fixed at positions $\mathbf{0}$ and $\mathbf{r}$ act as an external potential to the system of particle of type 2, namely

$$\hat{U}_{22}(\{\mathbf{r}_j\};\mathbf{0},\mathbf{r}) = U_{22}(\{\mathbf{r}_j\}) + \sum_{j=1}^{N_2}\phi_{12}(r_j) + \sum_{j=1}^{N_2}\phi_{12}(|\mathbf{r}-\mathbf{r}_j|). \tag{57}$$

Accordingly, Eq. (56) can be expressed as

$$\left\langle \exp\left[-\beta U_{12}^{(2)}(\mathbf{0},\mathbf{r})\right]\right\rangle_{z_2} = \frac{\hat{\Xi}_2(r;z_2,T,V)}{\Xi_2(z_2,T,V)}, \tag{58}$$

where $\hat{\Xi}_2(r;z_2,T,V)$ is the grand partition function of the system interacting with the Hamiltonian $\hat{U}_{22}(\{\mathbf{r}_j\};\mathbf{0},\mathbf{r})$, and it depends only on $r = |\mathbf{r}|$ as a result of the averaging. With this notation, Eq. (32) now takes the form

$$\omega_2(r;z_2,T) = -k_\mathrm{B}T\ln\left[\frac{\hat{\Xi}_2(r;z_2,T,V)}{\hat{\Xi}_2(r\to\infty;z_2,T,V)}\right], \tag{59}$$

implying

$$-\nabla_r\omega_2(r;z_2,T) = k_\mathrm{B}T\frac{\nabla_r\hat{\Xi}_2(r;z_2,T,V)}{\hat{\Xi}_2(r;z_2,T,V)}. \tag{60}$$

Using Eqs. (56) and (57), Eq. (60) takes the form

$$-\nabla_r\omega_2(r;z_2,T) = \left\langle\sum_{j=1}^{N_2}\left[-\nabla_r\phi_{12}(|\mathbf{r}-\mathbf{r}_j|)\right]\right\rangle_{\hat{U}_{22},z_2} = \left\langle\sum_{j=1}^{N_2}\mathbf{f}_{12}(|\mathbf{r}-\mathbf{r}_j|)\right\rangle_{\hat{U}_{22},z_2}, \tag{61}$$

where the notation $\langle\cdots\rangle_{\hat{U}_{22},z_2}$ denotes an expectation value carried in a system of particles 2 having activity z_2, and in which two particles of type 1 are *physically inserted* at positions $\mathbf{0}$ and $\mathbf{r}$. The quantity $\mathbf{f}_{12}(|\mathbf{r}-\mathbf{r}_j|) = -\nabla_r\phi_{12}(|\mathbf{r}-\mathbf{r}_j|)$ is the microscopic force acting

on particle 1 at position $\mathbf{r}$ due to a particle of type 2 at position $\mathbf{r}_j$. Denoting, in addition, as $\mathbf{f}_{11}(r) = -\nabla_r \phi_{11}(r)$ the microscopic force between particles of species 1, and taking into account Eqs. (38), (55), and (61), we obtain

$$\mathbf{F}_{\rm eff}(r; z_2, T) = \mathbf{f}_{11}(r) + \left\langle \sum_{j=1}^{N_2} \mathbf{f}_{12}(|\mathbf{r} - \mathbf{r}_j|) \right\rangle_{\hat{U}_{22}, z_2} . \tag{62}$$

Eq. (62) above is an extremely useful result and a very elegant one in its simplicity and clarity: it states that the total effective force being experienced by a particle of species 1 at position $\mathbf{r}$, is the sum of the direct force acting upon it by the other particle of species 1 held fixed at the origin, plus the *expectation value* of all the microscopic forces acting on it by all the particles of species 2 present in the mixture. We emphasize, once again, a crucial difference with the results we obtained in Section 2.2. There, all expectation values were taken at the Hamiltonian of a system of a pure reservoir of species 2-particles. In that case, the only physically active interactions were those between particles of the reservoir and the insertions of particles of species 1 were *virtual*: expectation values of observables involving 1-2 interactions were calculated in the sense of Widom insertions, which do not alter the configurations of particles of species 2. Here, instead, the two particles of species 1 are *physically* inserted and their simultaneous presence at positions $\mathbf{0}$ and $\mathbf{r}$ creates an inhomogeneous density profile $\rho_2(\mathbf{r}'; \mathbf{0}, \mathbf{r}, z_2, T)$ for the particles of species 2, the latter being the expectation value of the density operator $\hat{\rho}_2(\mathbf{r}'; \{\mathbf{r}_j\}) = \sum_{j=1}^{N_2} \delta(\mathbf{r}' - \mathbf{r}_j)$, viz:

$$\rho_2(\mathbf{r}'; \mathbf{0}, \mathbf{r}; z_2, T) = \left\langle \sum_{j=1}^{N_2} \delta(\mathbf{r}' - \mathbf{r}_j) \right\rangle_{\hat{U}_{22}, z_2} . \tag{63}$$

We can, thus, equivalently to Eq. (62) write

$$\mathbf{F}_{\rm eff}(r; z_2, T) = \mathbf{f}_{11}(r) + \int \mathbf{f}_{12}(|\mathbf{r} - \mathbf{r}'|) \rho_2(\mathbf{r}'; \mathbf{0}, \mathbf{r}, z_2, T)\, \mathrm{d}\mathbf{r}'. \tag{64}$$

Eq. (64) is very useful for theoretical approaches, in which one employs, e.g., density functional theory[30] or the superposition approximation[31] to calculate the density profile $\rho_2(\mathbf{r}'; \mathbf{0}, \mathbf{r}, z_2, T)$, from which the effective force results by performing (numerically) the integral of Eq. (64).

The *virtual insertion formalism* of Section 2.2 and the *physical insertion formalism* presented here are, of course, fully equivalent. Each one has its advantages and drawbacks, depending on what one wants to do. As previously mentioned, the virtual insertions are extremely inefficient in measuring directly the effective potential in a dense system with strong steric repulsions between the particles. Here, the physical insertion formalism is well-suited to deal with such problems. We just set up a MC or MD simulation, in which we first insert the two big particles into the empty box, and fix them at the desired positions, $\mathbf{0}$ and $\mathbf{r}$. Thereafter, the small particles are inserted sequentially, taking care of avoiding too strong overlaps. From then on, the system is equilibrated and the simulation, in which only particles of species 2 are moved, proceeds normally. At given intervals, the vector sum of the microscopic forces from all particles of species 2 on the big particle held fixed at position $\mathbf{r}$, $\sum_{j=1}^{N_2} \mathbf{f}_{12}(|\mathbf{r} - \mathbf{r}_j|)$, is measured and averaged at the end. A good check is to also measure the same forces $\sum_{j=1}^{N_2} \mathbf{f}_{12}(r_j)$ acting on the big particle at the origin, and

verify at the end of the run that the expectation values satisfy the *actio-reactio* law, as they should. Having gathered the statistics and calculated $\mathbf{F}_{\rm eff}(r; z_2, T)$ as given by Eq. (62), the effective potential $\phi_{\rm eff}(r; z_2, T)$ follows by simple, one-dimensional integration:

$$\phi_{\rm eff}(r; z_2, T) = \int_r^\infty |\mathbf{F}_{\rm eff}(r'; z_2, T)| \mathrm{d}r'. \tag{65}$$

3.2.2 Constrained systems

It will come as no surprise that for the case of constrained systems we obtain the same result as Eq. (62); still, it is instructive to go through the steps and demonstrate this explicitly. From Eq. (48), we obtain

$$\mathbf{F}_{\rm eff}(\mathbf{r}) \equiv -\nabla_{\mathbf{r}} \phi_{\rm eff}(\mathbf{r}) = -\nabla_{\mathbf{r}} F(\mathbf{r}). \tag{66}$$

Set explicitly $\mathbf{s}_{01} = \mathbf{0}$ and $\mathbf{s}_{02} = \mathbf{r}$ in Eq. (44) and define the constrained configuration partition function $\mathcal{W}_2(\mathbf{r})$ as

$$\begin{aligned} \mathcal{W}_2(\mathbf{r}) = & \int \mathrm{d}\mathbf{s}_{11} \cdots \int \mathrm{d}\mathbf{s}_{1M} \int \mathrm{d}\mathbf{s}_{11} \cdots \int \mathrm{d}\mathbf{s}_{2M} \\ & \times \exp\left[-\beta U_{11}(\mathbf{0}, \{\mathbf{s}_{i1}\})\right] \exp\left[-\beta U_{22}(\mathbf{r}, \{\mathbf{s}_{j2}\})\right] \\ & \times \exp\left[-\beta U_{12}(\mathbf{0}, \mathbf{r}, \{\mathbf{s}_{i1}\}, \{\mathbf{s}_{j2}\})\right] \\ & = \exp[-\beta F(\mathbf{r})], \end{aligned} \tag{67}$$

where $1 \leq i, j \leq M$. In analogy with the discussion in Section 3.2.1, define a modified interaction, $\hat{U}$, in which the coordinates $\mathbf{s}_{01}$ and $\mathbf{s}_{02}$ are clamped at the values $\mathbf{0}$ and $\mathbf{r}$, respectively:

$$\hat{U}(\{\mathbf{s}_{i1}\}, \{\mathbf{s}_{j2}\}; \mathbf{0}, \mathbf{r}) = U_{11}(\mathbf{0}, \{\mathbf{s}_{i1}\}) + U_{22}(\mathbf{r}, \{\mathbf{s}_{j2}\}) + U_{12}(\mathbf{0}, \mathbf{r}, \{\mathbf{s}_{i1}\}, \{\mathbf{s}_{j2}\}). \tag{68}$$

The instantaneous value of the microscopic force $\mathbf{f}_{\rm micro}(\mathbf{r})$ acting on the effective coordinate located at $\mathbf{s}_{02} = \mathbf{r}$ is derived from the interactions as

$$\mathbf{f}_{\rm micro}(\mathbf{r}) = -\nabla_{\mathbf{r}} U_{22}(\mathbf{r}, \{\mathbf{s}_{j2}\}) - \nabla_{\mathbf{r}} U_{12}(\mathbf{0}, \mathbf{r}, \{\mathbf{s}_{i1}\}, \{\mathbf{s}_{j2}\}), \tag{69}$$

i.e., it is the sum of all intramolecular and intermolecular microscopic forces. From Eqs. (66), (67), and (69) above, it immediately follows that the effective force is the expectation value of all microscopic forces taken in the Hamiltonian with interaction $\hat{U}$:

$$\mathbf{F}_{\rm eff}(r) = \langle \mathbf{f}_{\rm micro}(r) \rangle_{\hat{U}}, \tag{70}$$

where we set $\mathbf{r} \to r$ due to isotropy of space. Note that the microscopic force includes one term that is not fluctuating, and this is the direct interaction between the effective coordinate $\mathbf{s}_{01}$ and the coordinate $\mathbf{s}_{02}$; since they are clamped, the direct force between the two remains fixed. All other degrees of freedom fluctuate, and the averaging acts on those microscopic forces. This completes the analogy with the discussion in Section 3.2.1 and in particular with Eq. (62). For more details and applications of this method to polymeric systems, we refer the reader to Refs.[32,33]

3.3 Generalized Widom insertion

A different possiblity to compute the effective interaction between bonded and intrinsically fluctuating macromolecular aggregates is a generalization of the Widom insertion algorithm.[34,22,35] Indeed, combining Eqs. (43), (47), (48), and (67), we obtain

$$\exp\left[-\beta\phi_{\mathrm{eff}}(\mathbf{r})\right] = \frac{\mathcal{W}_2(\mathbf{r})}{\mathcal{W}_1^2}. \tag{71}$$

By virtue of Eq. (68), we readily obtain

$$\frac{\mathcal{W}_2(\mathbf{r})}{\mathcal{W}_1^2} = \left\langle \exp\left[-\beta U_{12}(\mathbf{0},\mathbf{r},\{\mathbf{s}_{i1}\},\{\mathbf{s}_{j2}\})\right]\right\rangle_{U_{\mathrm{intra}}}, \tag{72}$$

where the notation $\langle\cdots\rangle_{U_{\mathrm{intra}}}$ indicates that the expectation value has to be calculated in the ensemble of the Hamiltonian whose interaction part contains macromolecules that do not interact with one another; this is equivalent, of course, to just a single macromolecule. Finally, Eqs. (71) and (72) yield

$$\phi_{\mathrm{eff}}(\mathbf{r}) = -k_{\mathrm{B}}T \ln \left\langle \exp\left[-\beta U_{12}(\mathbf{0},\mathbf{r},\{\mathbf{s}_{i1}\},\{\mathbf{s}_{j2}\})\right]\right\rangle_{U_{\mathrm{intra}}}. \tag{73}$$

It might appear at first sight paradoxical that one is capable of expressing a constrained free energy as the expectation value of some quantity. However, as Eq. (48) readily shows, $\phi_{\mathrm{eff}}(\mathbf{r})$ is a *difference* between two constrained free energies, one at separation $\mathbf{r}$ and the other at infinite separation. Free energy differences can indeed be calculated very efficiently in computer simulations.

Widom insertion[22,35] takes advantage of Eq. (73) in the following way. First, a very large number of independent and equilibrated single-molecule configurations are generated. Thereafter, these are combined in pairs by simply pulling one of the two is such a way that the effective coordinates lie at a distance $r = |\mathbf{r}|$ from one other, i.e., one molecule is inserted at a distance r from the other. From the ensemble of these inserted pairs, the expectation value appearing in Eq. (73) is computed. The method is simple and transparent; however, it is inefficient when the typical conformations of the interacting entities are markedly different from those generated within the non-interacting Hamiltonian. For instance, Widom insertion would not be appropriate to calculate the effective interaction of a dense polymer brush with a hard wall, since for close brush-wall approaches, the massive retraction of the brush hairs, enforced by the presence of the wall, would result in configurations that appear extremely rarely in the free-brush case. However, Widom insertion is well-suited for fractal, open, and penetrable macromolecules, whose effective interactions do not exceed a few $k_{\mathrm{B}}T$ even at the closest approaches. It has been successfully employed to linear chains,[22] sparsely polyelectrolyte-coated colloids and dendrimers,[35] as well as ring polymers,[36] for which special care must be taken to preserve the topological constraint of no-concatenation.

4 Hard-Soft Colloidal Mixtures

Highly asymmetric binary mixtures have long been investigated as model systems for colloid-polymer suspensions.[37,10,1,16] Indeed, when the polymer-to-colloid size ratio and polymer density are sufficiently small one can think of tracing out the microscopic degrees

of freedom of the polymer and thus obtain an effective isotropic potential for the polymer-polymer and polymer-colloid interactions. Performing a further coarse-graining on the small-component remaining degrees of freedom allows for the calculation of an effective colloid-colloid potential. The tunability of this *depletion* interaction, which depends on the polymer (also called *depletant*) nature, size and density, makes it possible to finely control the degree of association between large particles in solution. This is of great importance for both soft matter and biological systems.[1,38]

In the approximation of hard polymer-colloid interaction and ideal (i.e. zero) polymer-polymer interactions the effective colloid-colloid potential can be derived analytically.[13,14] This Asakura-Oosawa-Vrij formulation has become the reference model system for depletion interactions and its use is widespread.[15,10,39–41,16] One can go one step further and calculate the effective interaction also for the case of non-ideal depletants interacting through a hard-sphere repulsion.[42] However, it has been shown that both polymer-polymer and polymer-colloid effective potentials always show some degree of softness,[34] which makes it extremely hard, if not impossible, to analytically calculate Eq. (19). It is thus more convenient to provide a numerical estimate of the effective interaction between the colloids. Given the extremely simple nature of the components, which interact only *via* isotropic potentials, the best approach is the one outlined in Section 3.1: the inversion of the radial distribution function $g(r)$ in the limit of infinite dilution, *i.e.* for $\rho \to 0$. In practice, this zero-density limit is reached when three-body effects are negligible[c]. Care has to be taken in choosing the right simulation parameters.

In the following, σ is the colloid diameter, q is the depletant-to-colloid aspect ratio and $\sigma_d = q\sigma$ and ρ_d are the depletant diameter and number density, respectively. For all but the simplest cases, computing the $g(r)$ with the required numerical accuracy in the whole r-range of interest is not as straightforward as it may seem at a first glance. Indeed, the most simple approach, *i.e.* performing an unbiased simulation of a very diluted suspension of colloids in a sea of small depletants, has several practical drawbacks, namely:

1. The number of depletants N_d required to simulate the mixture becomes prohibitively large for small values of q. For example, computing the radial distribution function $g(r)$ of 100 colloids at a density $\rho\sigma^3 = 0.01$, with $q = 0.1$ and $\rho_d(q\sigma)^3 = 0.2$, would require to simulate no fewer than 10^7 particles!

2. Since the number of depletants is proportional to the number of colloids, one could keep N_d low by simulating just two colloids. However, in doing so, the numerical accuracy of the resulting $g(r)$, which depends on the number of pairs of colloids, would be severely hindered.

3. The probability $P(r)$ of finding two colloids at a distance comprised between r and $r + \mathrm{d}r$ is connected to the $g(r)$ *via* the relation

$$P(r) = P_{\mathrm{id}}(r) \cdot g(r), \tag{74}$$

[c]This is not to say that having just two colloids in a sea of depletants is a sufficient condition to enforce the $\rho \to 0$ limit. Indeed, if the length of the simulation box side is comparable with the extent of the distortions induced by the presence of the colloids on the density profile of the depletants, a colloid could feel twice the presence of the second colloid thanks to periodic boundary conditions.

where $P_{\mathrm{id}}(r)$ is the corresponding quantity for an ideal gas. In three dimensions, $P_{\mathrm{id}}(r) = 4\pi\rho_d r^2$. Since, for sufficiently large r, $g(r) \simeq 1$ and hence $P(r) \sim r^2$, it follows that colloids will more likely be far apart from each other, wasting compute time on configurations which are, sampling-wise, uninteresting.

4. As mentioned in Section 3.1, for even the most simple systems the $g(r)$ will never be sampled uniformly: distances corresponding to peaks in the $g(r)$, which translate to local minima in $\phi_{\mathrm{eff}}(r)$, will be visited with a larger probability than distances which correspond to maxima of the effective potential. On top of this, very attractive systems will end up trapping colloids in local minima, thus further hindering the sampling.

A simple, albeit both general and powerful, way of achieving a quasi-uniform sampling is to use the 'window sampling' presented in Section 3.1. In summary, the interval of interest $[r_{\min}, r_{\max}]$ is split into N_w windows. In each window j, centered around its midpoint r_j, the total potential felt by the two large particles has an additional term, $\Phi_{\mathrm{bias}}(r; j)$. Splitting the region of interest into different windows has several advantages:

1. All the windows can be sampled separately, making the overall process inherently parallel.

2. The interval $[r_{\min}, r_{\max}]$ is not fixed once and for all: if need be, more windows can be added.

3. The width of each window, Δr_j, can be chosen as to yield an as uniform sampling as possible. For example, if we observe that the $g(r)$ within a given window varies substantially, we can decide to split the window up in order to achieve a better sampling.

4. We can control the convergence of the signal window by window. We can thus decide to increase the length of the simulation only for those windows having unsatisfactory statistics.

The choice of the local bias potential is not to be taken lightly, as different $\Phi_{\mathrm{bias}}(r; j)$ can have a different impact on the convergence, the numerical accuracy of the final $g(r)$ and the required number of windows. The two most common choices are a harmonic potential and an infinite well. The former, also known as umbrella sampling,[23] is most useful when performing molecular dynamics simulations and when large particles have internal degrees of freedom, *e.g.* star polymers. The infinite well bias is more suited when dealing with simple objects, such as spheres. This local bias is given by

$$\Phi_{\mathrm{bias}}(r; j) = \begin{cases} \infty & \text{if } r < r_j - \frac{\Delta r_j}{2} \\ 0 & \text{if } r \in \left[r_j - \frac{\Delta r_j}{2}, r_j + \frac{\Delta r_j}{2}\right] \\ \infty & \text{if } r > r_j + \frac{\Delta r_j}{2}, \end{cases} \tag{75}$$

where Δr_j is the width of the window. Neighboring windows j and $j+1$ should be overlapping, *i.e.* $r_j + \frac{\Delta r_j}{2} > r_{j+1} - \frac{\Delta r_{j+1}}{2}$. The bias potential (75) may be used only in conjunction with Monte Carlo simulations, since it is not derivable at the window boundaries. In what follows we will show how to use this infinite well potential to extract the

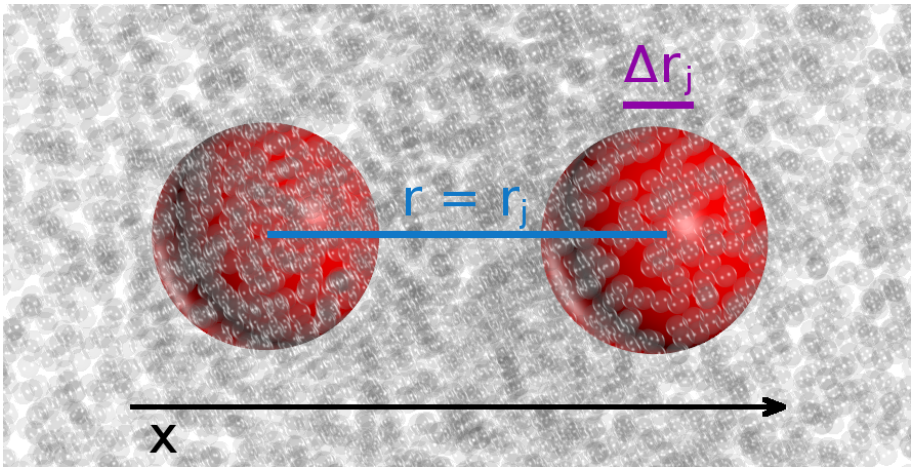


Figure 2: A cartoon exemplifying the sampling procedure. It shows two colloids immersed in a sea of small polymers of size ratio $q = 0.1$. The black line is the direction of the x axis, along which both colloids lie. The blue and violet segments are the distance r between the two colloids, set to the midpoint of the window r_j, and the maximum allowed displacement Δr_j, respectively.

effective interaction between the large components of a mixture of soft colloids and polymers. The softness in the colloid-colloid and colloid-polymer interactions is provided by power-law potentials:

$$V_{36}(r) = \epsilon \left(\frac{\sigma_{ij}}{r}\right)^{36}, \tag{76}$$

where ϵ sets the energy scale and σ_{ij} is the diameter associated to the specific interaction ij. $\sigma_{cc} = \sigma$ and $\sigma_{cp} = \sigma(1+q)/2$ for colloid-colloid and colloid-polymer interactions, respectively. We model the polymers as weakly repulsive gaussian particles,[43] with the polymer-polymer interaction potential given by

$$V_{pp}(r) = 2k_{\mathrm{B}}T \exp\left[-\frac{r^2}{(0.6q\sigma)^2}\right]. \tag{77}$$

We start off by noting that, without loss of generality, we can use Eq. (74) to our advantage by constraining the two large colloids to lie on the x axis, so that $g(r) = P(r)/\rho_d$. This expedient simplifies the data analysis and may lead to a higher numerical accuracy when r is small, since we avoid dividing by r^2. Figure 2 shows a cartoon depicting the geometry and the employed method for a generic window j. For this specific test case we set $\rho_d\sigma_d^3 = 0.25$, $N_w = 21$, $\Delta r_1 = 0.02\sigma$, $\Delta r_{j>1} = 0.024\sigma$, $r_{\min} = 0.92\sigma$ and $r_{\max} = 1.34\sigma$. The overlap between neighboring windows is set to 0.004σ.

We start a simulation in each window j by placing the two colloids in their initial positions, $\mathbf{r}_1 = (0,0,0)$ and $\mathbf{r}_2 = (r_j,0,0)$, and the polymers in random positions, only taking care that no overlap with the colloids is generated. Then we run a regular Monte Carlo simulation with translational moves, with the constraints that

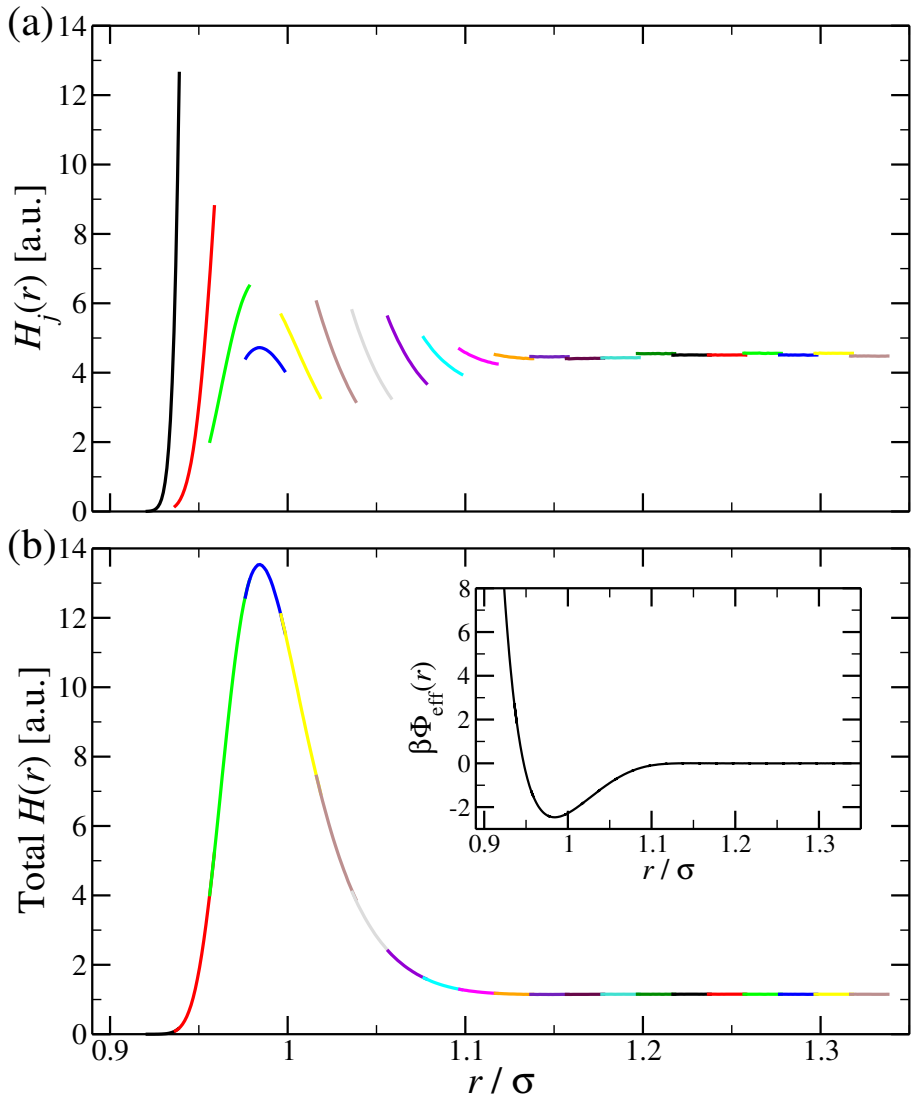


Figure 3: (a) Colloid-colloid distance histograms for all the investigated windows. The scale on the y axis is arbitrary. (b) Total $H(r)$: all the histograms in (a) have been rescaled so as to superimpose the regions of overlap between the windows. This curve is proportional to the radial distribution function $g(r)$. Inset: the effective colloid-colloid potential $\phi_{\rm eff}(r)$, calculated inverting the $H(r)$ and then shifting it up so that $\lim_{r\to\infty}\phi_{\rm eff}(r) = 0$.

(i) trial displacements for colloids are parallel to the x axis and (ii) trial moves for which $|\mathbf{r}_2 - \mathbf{r}_1| \notin \left[r_j - \frac{\Delta r_j}{2}, r_j + \frac{\Delta r_j}{2}\right]$ are outright rejected. We equilibrate the system for 10^6 MC sweeps, each sweep being equivalent to $N = N_d + 2$ trial moves, and then start accumulating data by updating the histograms of the colloid-colloid distances, $H_j(r)$.

Figure 3(a) shows the output of a typical set of simulations, with a curve for each window.

Once we have collected all the histograms, we have to stitch them together in order to obtain a function which is proportional to the $g(r)$. The process begins from the first window and then joins subsequent windows by applying a least square method to each overlap region between neighbouring windows j and $j+1$. In practice, adding $H_{j+1}(r)$ to the total $H(r)$ is done by rescaling the former by a factor

$$c_{j,j+1} = \frac{\sum_k r_{k;j}^2}{\sum_k r_{k;j} r_{k;j+1}}, \tag{78}$$

where k runs over all the points which are in the overlap region between the two windows. The result of this operation is shown in Figure 3(b). The output quantity of this procedure, namely the total $H(r)$, is proportional to the $g(r)$ of the system. One can either set the unknown proportionality constant by imposing the condition $\lim_{r\to\infty} g(r) = 1$ or, equivalently, by using Eq. (52) and then imposing $\lim_{r\to\infty} \phi_{\text{eff}}(r) = 0$. In this framework, the $r \to \infty$ limit is effectively reached when the $g(r)$ reaches its final value at large values of r. If the $g(r)$ does not display any plateau at large distances, more windows should be added in order to reach the asymptotic behavior.

Finally, we note that Eq. (19) can be rewritten as:[10]

$$\exp\left[-\beta\Omega(\{\mathbf{R}_i\}; N_1, z_2, V, T)\right] = \exp\left[z_2 \int \mathrm{d}\mathbf{r} \exp\left[-\beta(U_{12} + U_{22})\right]\right]. \tag{79}$$

If the depletants are ideal among themselves, from Eq. (79) follows that $\Omega(\{\mathbf{R}_i\}; N_1, z_2, V, T)$, and hence the effective pair potential, depends linearly on the depletant density $\rho_d = z_2$, which plays the role of an inverse temperature. In other words, once we compute the effective potential at a given density ρ_d^*, $\phi_{\text{eff}}(r; \rho_d^*)$, it immediately follows that the effective potential at a different depletant density can be obtained as

$$\phi_{\text{eff}}(r; \rho_d) = \frac{\rho_d}{\rho_d^*} \phi_{\text{eff}}(r; \rho_d^*). \tag{80}$$

Unfortunately, there are no available short cuts in the case of non-ideal depletants: we have to compute as many effective potentials as depletant densities we need to investigate.

The effective potential, shown in the inset of Figure 3(b), can now be used to simulate the one-component system and extract colloid-related structural and dynamical data. As an example, we simulate a one-component system made of 10000 colloids interacting through the effective potential and make a comparison with the full binary mixture (100 colloids and 250000 depletants). For this specific case, the simulation of the full binary mixture required a few days of computing time on a GPU. Simulating the same state point as a one-component system, on the other hand, took only a few hours on a single core. This is roughly equivalent to a 10^4-fold performance increase.

Figure 4 shows the structure factors $S(q)$ of the two systems.[27] The two curves fall on top of each other within statistical accuracy. This result demonstrates that the choice of coarse-graining the interactions to obtain an effective pair-potential is justified: the excellent agreement between the one-component and full mixture results shows that many-body effects are negligible even for moderately high ρ_d values such as the one investigated here.

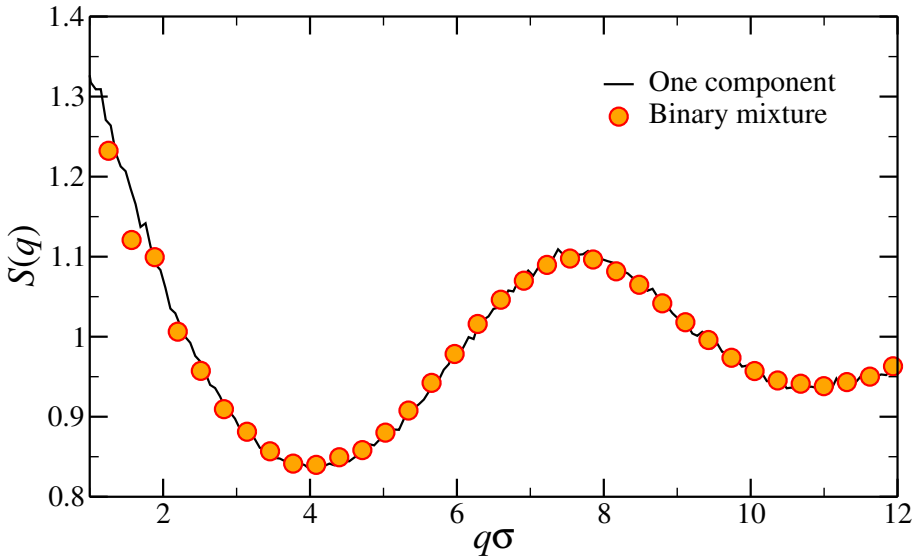


Figure 4: Structure factor $S(q)$ of the one-component system (black line), where colloids interact through the computed effective potential $\phi_{\text{eff}}(r)$, and of the full binary mixture (orange points).

5 Dendrimers

In considering polymers, one usually thinks about long chains of repeating units. However, in the late 1970's an interest arose in r egular and highly branched structures that resulted in the first synthesis of dendrimers by Vögtle.[44] A coarse-grained representation of such a macromolecule is provided in Figure 5. Similar to polymer chains there is a simple unit-structure that is repeated, but in the case of the dendrimer shown here it is chosen to be trifunctional and results in a network rather than a linear chain. The center or seed of the dendrimer is formed by a connected pair of the so-called generation 0 monomers and consecutive generation layers can be added by for instance a two step process. First the non-bonded ends need to be activated or unblocked by some agent which is removed from the environment afterwards. Now an excess of trifunctional units is added to the system and each of the four reactive ends in the seed can bind a single unit. Given enough time for all sites to react and removing the remaining, non-reacted trifunctional units, one is left with macromolecules that each consist of six units and of which only the outer four have two (eight in total) blocked ends. Repeating the same procedure allows one to add eight more units in the second generation shell, sixteen in the third, and so on. This results in macromolecules that will contain an exponentially growing number of units as function of the number of generation or iteration steps.

Since their first appearance, these type of macromolecules have received a considerable interest from both experimental and theoretical research.[45] This is not only due to their role in applications such as solubility enhancement, drug-delivery vectors, or nano-carriers[46–49] but also from a fundamental point of view because they combine properties of colloidal and polymeric systems.[50]

If we consider the interaction between two uncharged dendrimers we find an effective pair-potential with a Gaussian shape.[51,52] Given the overall spherical shape, where the number of constituents grows exponential but, due to connectivity, the radius grows only linearly with increasing generation, one can control the softness from penetrable, low generation to almost hard-core, high generation dendrimers. The richness in behavior is increased further by allowing for different units, for instance by adding charge to the outer generation monomers but keeping neutral units in the center, which results in spherical particles that in a first order approximation have a homogeneous surface charge density.[53] Accordingly, dendrimers[54–57] play a prominent role in soft matter research.[58–65]

A significant amount of research has been devoted to properties of neutral dendrimers, with the focus on internal structure, interactions, and phase behavior as function of the concentration of the solution.[66–69] Polyelectrolyte dendrimers are less well studied, for the obvious reason that the addition of long range Coulomb interaction increase the complexity of the system. A good example is the fact that in a water-based solution of such dendrimers, the ionizable groups dissociate which results in charged macromolecules and oppositely charged counterions that can move freely in solution. Since the counterions are only weakly bound to the dendrimer, the internal structure responds to the internal charge distribution and this causes the molecule to swell.[70,71] The interactions and behavior can be influenced and controlled further by screening the Coulomb interaction via the addition of salt, manipulating the charge distribution through a modification of pH, or adjusting the spacer length between the nodes in the dendrimer structure.[72–75]

5.1 Simulation Model

An appropriate method to study polyelectrolyte dendrimers is by means of computer simulations, in particular monomer-resolved Molecular Dynamics simulations that are com-

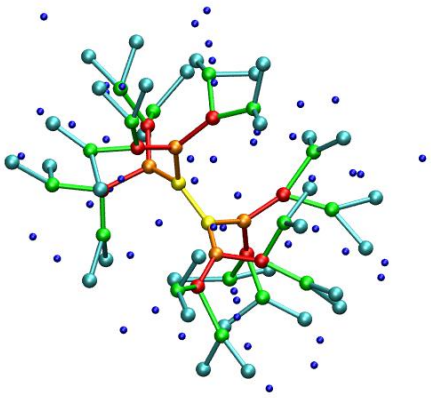


Figure 5: A coarse-grained representation of a polyelectrolyte dendrimer of generation 4 and functionality 3 with a seed formed by a pair of particles (yellow). Higher generation shells consists of 4 (orange), 8 (red), 16 (green), and 32 (cyan) units. In order to satisfy charge neutrality, additional counterions (blue) have been added to the system.

plemented with the Ewald summation technique to treat the electrostatic interactions.[76,77] In order to concentrate on a qualitative behavioral description of these complex systems, we employ a coarse-grained model in which the atomistic structure and covalent bonds are replaced by a simple bead-spring model.

For the center of the dendrimer a single, bonded pair of monomeric units is chosen, which we can label as the two 0th generation building blocks. By assuming the usage of trifunctional units only, every particle in the most outer shell has two unused reaction sites. Hence, on augmenting the generation of the dendrimer, the number of particles in the next shell is twice as large, i.e., the gth shell of the dendrimer consists of 2^{g+1} units. If we denote by G the generation of the most outer shell, we find by adding all shells that such a dendrimer of generation G contains a total of $2^{G+2}-2$ particles. In our case, we restrict ourselves to dendrimers of generation 4 and therefore will contain 62 units or nodes. We can modify the spacer length P (in terms of bondlengths), by inserting a short chain of $P-1$ bifunctional units between the nodes. In our case, this means that the total number of units (bi- and trifunctional) is given by $61P+1$, which follows from the realization that is we start with a single unit the addition of a node always increases the number of spacer chains by one as well.

By carefully selecting suitable constituents for their synthesis, dendrimers can be obtained that are charge neutral at large pH, but of which the nodes become ionized at low pH.[50] During the synthesis process one is of course not obliged to use the same kind of building blocks at every reaction step. With the aid of functional groups, such as secondary and tertiary amines, one can obtain structures where the monomers in outer shells are already ionized for intermediate pH values. It is therefore possible to influence the charge contribution in polyelectrolyte dendrimers in a controlled fashion. Here we will focus on three different charge distributions, with ionizable groups located only at the nodes. These are the fully neutral case, only charged end-groups, and fully charged dendrimers, with an overall charge number of $Z=0$, $Z=32$, and $Z=62$ respectively. For computational reasons the spacers are chosen to be neutral and have a fixed length throughout the dendrimer that is restricted to a maximum of 5 ($P=1$ up to $P=6$).

In the bead-spring model we employ here for athermal solvent conditions, all units, that are either nodes in the dendrimer or located within the inter-node chains, are represented by identical spherically symmetric monomers. The mutual interaction between any pair these monomers is modeled via the purely repulsive, truncated and shifted Lennard-Jones[78] potential that only depends on the relative distance r

$$V_{\mathrm{LJ}}(r)=\begin{cases}4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12}-\left(\frac{\sigma}{r}\right)^{6}+\frac{1}{4}\right] & r\leq 2^{1/6}\sigma\\ 0 & r>2^{1/6}\sigma,\end{cases} \tag{81}$$

where ϵ represents the strength of the interaction and σ is a measure for the diameter of the particles. It is convenient to use ϵ and σ as the units of energy and length respectively. The chemical bonds between monomers that determine the connectivity in the dendrimer are typically covalent in nature and can be described by the so-called finite extensible nonlinear elastic (FENE) potential[79]

$$V_{\mathrm{FENE}}(r)=\begin{cases}-U_0\left(\frac{R_0}{\sigma}\right)^2\ln\left[1-\left(\frac{r}{R_0}\right)^2\right] & r\leq R_0\\ \infty & r>R_0.\end{cases} \tag{82}$$

Note that the potential diverges at both distances $r = 0$ and $r = R_0$, which provides the minimum and maximum bond length respectively. The strength of this spring is measured by means of U_0. The particular values in the current model are chosen to be $U_0 = 15\epsilon$ and $R_0 = 1.5\sigma$. This choice ensures that there will not be any unphysical crossing of bonds within the simulations, because the energy barrier that would be needed to overcome lies outside the range of the typical energy fluctuations found in these systems.[80]

We have implemented the Ewald summation method to simulate the long range Coulomb interactions between charged particles that is given by

$$V_{\mathrm{Coulomb}}(r) = k_{\mathrm{B}} T \lambda_{\mathrm{B}} \frac{Z_i Z_j}{r} \tag{83}$$

with Z_i and Z_j the charge numbers, k_{B} the Boltzmann constant, T the temperature and λ_B the Bjerrum length

$$\lambda_{\mathrm{B}} = \frac{e^2}{\epsilon_{\mathrm{r}} k_{\mathrm{B}} T}, \tag{84}$$

where ϵ_{r} is the relative permittivity. The charged dendrimers that we consider carry monovalent charges which are located on all the 62 nodes or on the 32 terminal monomers only. Their electrostatic interaction is added on top of the always present Lennard-Jones and FENE potentials. In order to guarantee charge neutrality of the full system 62 or 32 oppositely charged monovalent counterions respectively have to be added per dendrimer. In addition, various salt concentrations of positive/negative salt ion-pairs have been added, where both components are either monovalent or divalent in nature. The three representative salt concentrations that will be used are 0.05, 0.10, and $0.50\ \mathrm{mol/l}$. In addition to their charge, all counter and salt ions interact sterically with all other particles by means of the same shifted Lennard-Jones potential (81) used for the monomers within the dendrimers.

The solvent of the solution is chosen to be water at room temperature. The water molecules are not simulated explicitly, but only implicitly by imposing $\epsilon_{\mathrm{r}} = 80$ and for which an interaction strength $k_{\mathrm{B}} T = 1.2\epsilon$ is chosen. The temperature is controlled via an Andersen thermostat. The corresponding Bjerrum length $\lambda_B = 0.714\,\mathrm{nm}$ is fixed to be $\lambda_{\mathrm{B}} = 3\sigma$, which results in a monomer diameter $\sigma = 0.238\,\mathrm{nm}$ typical for the van-der-Waals radii of atoms that can be used in the framework of a bead-spring model for polymer molecules.[1,81]

The shape of simulation box is cubic with periodic boundary conditions and sides of length $L \cong 11 R_g$, where R_g denotes the radius of gyration of a single isolated dendrimer. This quantity is a measure for the radius of extended objects, that in the case of $P = 1$ and $P = 2$ dendrimers in our model has typical values of 2.86σ or 4.55σ, respectively. In order to measure the effective interaction between dendrimers a single pair of dendrimers, the corresponding counterions, and salt ions are placed in the simulation box. Rather than measuring the effective interaction potential, $V_{\mathrm{eff}}(D)$, as a function of the distance $D = |\mathbf{R_1} - \mathbf{R_2}|$ between the centers of mass, we consider here the effective force at fixed distances, which can be measured directly in MD-simulations. To this end we fix the centers of mass $\mathbf{R}_1$ and $\mathbf{R}_2$ of the two dendrimers and maintain their location by external force fields. A convenient choice is to place both centers of mass on the diagonal of the simulation box, because this will minimize finite size effects arising from the periodic boundary conditions.

The relation between the effective force and potential (66) is given by

$$\mathbf{F}^{\alpha}_{\rm eff}(\mathbf{D}) = -\nabla_{\mathbf{R}_\alpha}\phi_{\rm eff}(|\mathbf{R_1} - \mathbf{R_2}|), \tag{85}$$

where the index $\alpha = 1, 2$ corresponds to either dendrimer. The effective interaction can be measured directly in the simulation, because it is the ensemble average of the total force acting on a dendrimer that follows from the addition of all forces $\mathbf{F}^{\alpha}_{i}$ acting on monomers within the same molecule, i.e.,

$$\mathbf{F}^{\alpha}_{\rm eff} = \left\langle \sum_{i=1}^{N_\alpha} \mathbf{F}^{\alpha}_{i} \right\rangle . \tag{86}$$

Note that the total instantaneous force has the same magnitude but opposite direction of the constraining force that fixes the center of mass location. Although the instantaneous forces acting on both dendrimers can be different, their ensemble averages should be the same. In addition, for reasons of symmetry, the effective force should be along the relative distance. Deviations from either of these observations can be used as a measure to determine whether proper equilibrated values have been obtained.

The time is measured in units of $\tau = \sqrt{m\sigma^2/\epsilon}$ and the time-step used in solving the equations of motion is $\delta t = 0.002\tau$. Initially, the dendrimers are placed with random configurations at their designated positions, and all ions are inserted at random locations throughout the simulation box. After an equilibration time that is long enough to relax all stresses within and between the dendrimers, as well as for the ions to diffuse in to the dendrimers, equilibrium measurements are made over a time of typically 10^7 time-steps.

5.2 Neutral Dendrimers

In these complex systems we can investigate in detail the consequences of particular modifications in the architecture on the molecular level. In particular, the effects that changes in generation number, spacer length, and charge distribution have on the effective interactions. In the case of neutral dendrimers in athermal solvents, one finds that the effective interaction potential in general has a Gaussian-like shape.[51,52] Its strength and range could be increased by for instance augmenting the generation number and introducing longer spacers respectively. In simulations by Götze *et. al.*,[66] the spacer length was modified by adjusting the softness of the interaction between the nodes of the dendrimers. Rather than changing the monomer-monomer interaction, this effect can be reproduced by means of the addition of spacer monomers.[53] It turns out that the radius of gyration R_g for fixed generation number scales with the spacer length P as $R_g \propto P^{3/5}$. In addition, the density profiles, which are a measure for the internal structure of the molecule, can be scaled to collapse onto a master-curve. Here, we will examine whether a similar behavior can be found for the effective interactions

Hereto we have performed simulations of two generation 4 dendrimers with spacer lengths $P = 1 - 6$ and measured the effective force at fixed relative distances. Typical snapshots of such a pair of neutral dendrimers interacting with each other are shown in Figure 6, where we let the centers of mass of both dendrimers coincide, i.e., $D = 0$. Since the mutual interaction has only steric components, it is not surprising that this limit corresponds to the maximum interaction. Rather than truly creating overlapping configurations,

where the two dendrimers penetrate each other, one mainly observes a segregation in different regions for either dendrimer, i.e., the dendrimers wrap around each other. This effect is strongest in the case of $P = 1$ with the shortest spacerlength. On increasing the spacerlength the overall size (radius of gyration) increases and the added flexibility allows more free volume to be explored by the individual monomers. Consequently a weaker overall interaction between the dendrimers is observed.

The radial component of the effective force between the dendrimers as a function of the relative distance D is shown in Figure 7(a) for four different values of P. On approaching each other from the far distance, non-interacting range one observes a slow increase in the repulsive interaction between the macromolecules. When the dendrimers get closer, more intermolecular monomer pairs add their repulsive contribution to the total effective force. What might be surprising, however, is that the repulsive force attains a maximum value at a relative distance close to the radius of gyration. When the centers of mass of the dendrimers are forced closer together, the repulsive force diminishes and ultimately vanishes at zero distance. This effect is due to the fact that for such configurations a substantial number of monomers become surrounded by other monomers from either the same or the other dendrimer, and hence the forces that such a monomer experiences start to cancel out. Therefore these units contribute less to the total repulsive force between the dendrimers. In addition, some of the outer monomers that experience typically a force directed away from the center of mass of the dendrimer they belong to, will, when projected on the line connecting the centers of both, lie beyond the center of the other dendrimer and hence give an attractive contribution to the effective interaction. The results that the force at zero distance should vanish, follows from the realization that the effective interaction is an ensemble average. Since in such configurations there should therefore be no preferred direction, the average effective force needs to be disappear. Note that this does not imply that in a particular configuration there can not be an either attractive or repulsive force between the dendrimers.

The measured effective forces for different spacer lengths P are very similar in shape and when scaled appropriately will collapse on a mastercurve as shown in the inset of Figure 7(a). This collapse is obtained when the force is scaled with a factor $P^{-0.65}$, where the scaling factor is fitted from the data presented here. In addition the distance D is measured in units of the radius of gyration R_g measured for a single, isolated dendrimer. It should be noted that the radius of gyration of the dendrimers on interacting is affected. On decreasing the relative distance, the radius of gyration increases with respect to that of an isolated dendrimer, obtaining a maximum value at zero center-to-center distance that is approximately 15% larger than its original size. This increase in dimensions is a direct

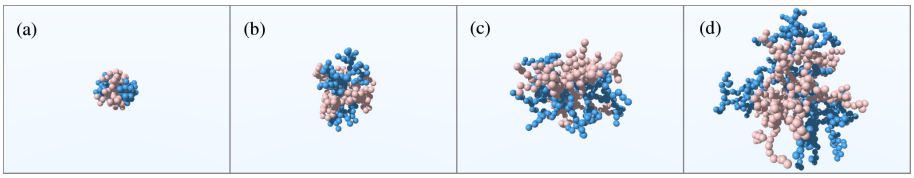


Figure 6: Snapshots from simulations of neutral dendrimers with (a) $P = 1$, (b) $P = 2$, (c) $P = 4$, and (d) $P = 6$ spacer segments. The distance between the centers of mass is $D = 0$ in all four cases.

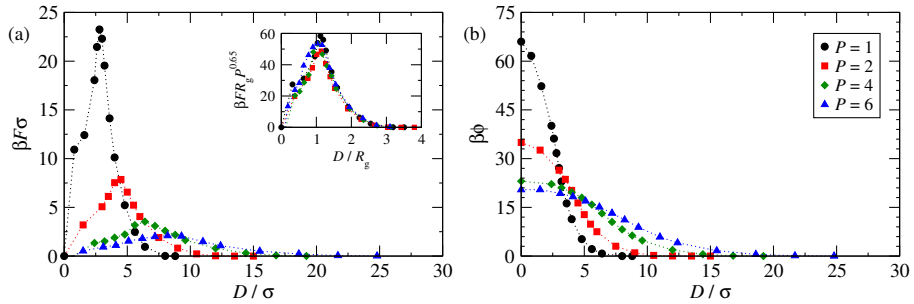


Figure 7: (a) Effective force F and (b) effective potential ϕ, between two neutral dendrimers with number of spacer segments P plotted against the distance D between their centers of mass. The inset illustrates that the data collapse when scaled appropriately.

consequence of the fact that the mutual overlap results in an increased local monomer density and a corresponding higher internal pressure. The larger size of the dendrimer that follows is achieved by a combination of stretching bondlengths and increasing the effective persistence length in the polymeric structure.

The integration of the effective forces with respect to the distance is straightforward and results in the effective interaction potentials shown in Figure 7(b), that have the characteristic Gaussian-like shape. The maximum that we observed in the force, corresponds to the inflection point in the potential where the concave switches to the convex behavior. The maximum at zero separation distance are a measure for the energy cost of overlapping dendrimers, which not surprisingly is highest for the smallest spacer length. This was already anticipated from the snapshots of Figure 6 that reveal a more open structure on augmenting the spacerlength thereby facilitating such configurations with respect to that of the compact structure at $P = 1$.

5.3 Charged Dendrimers

If proper functional groups are incorporated during the synthesis procedure of dendrimers, it is possible to influence the interaction between these macromolecules. In particular, by employing groups that get protonated at low pH values, allow for an easy control of the internal charge distribution that range from neutral to fully charge objects. Since we consider an implicit solvent here, we mimic the effect of pH by adding charges to particular chosen node monomers. To simplify the study of such charged structures, we restrict ourselves to three cases only. The fully neutral dendrimer already discussed, a dendrimer where every node carries a charge, and an intermediate structure where only the most outer nodes are charged. In the case of our 4th generation dendrimers this means that they have a total valency $Z = 0$, $Z = 62$, and $Z = 32$ respectively. Without loss of generality, we assume the charges on a node will be a positive elementary charge. In order to maintain an overall neutral system, negatively charged counter ions have to be added to the system.

In Figure 8 the effect on the conformations of dendrimers caused by changes in the pH or charging the macromolecules, is illustrated by representative snapshots. The figure shows configurations of a pair of these particles with spacer length $P = 4$ for each of the three charge distributions and two different relative distances, i.e., the overlapping case of

zero distance, and a distance equal to that of the radius of gyration where there is only a partial overlap. From the snapshots one can already infer that increasing the charge results in a monotonic swelling of the dendrimers, which is simply due to the internal electrostatic repulsion between charged monomers. The counter ions can move freely in and out of the dendrimer and are only loosely bound to the molecule. Only a small fraction, that depends on the overall density, will leave the vicinity of either molecule. The presence of counterions within the complex will, at least partially, screen the Coulomb interactions between both dendrimers.[74]

The electrostatic interactions not only result in a significant swelling for both types of charged molecules, but it also results in more intertwined rather than interpenetrating conformation if the two dendrimers approach each other. In the intermediate charged case, the stretching is less compared to the fully charged molecules. This is not only due to the weaker overall charge, but also caused by the fact that the core by itself will be neutral

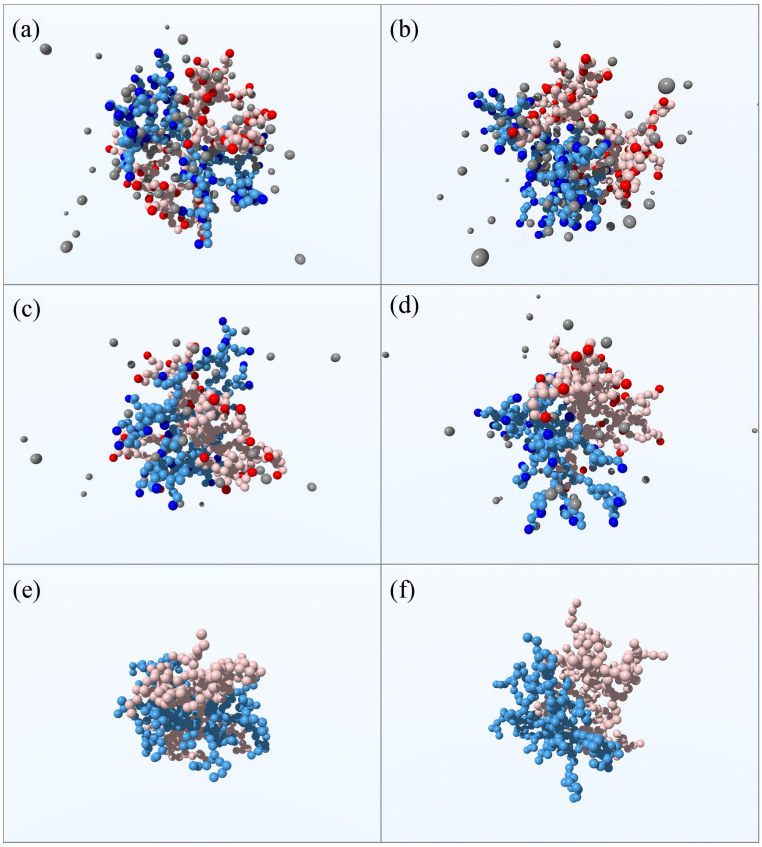


Figure 8: Snapshots from simulations of dendrimers with $G = 4$ and $P = 4$ at distance $D = 0\sigma$ for different pH values of the solvent: (a),(b) fully charged dendrimers, (c),(d) dendrimers with end-groups charged, and (e),(f) neutral dendrimers. The dendrimer's centers of mass either coincide, (a),(c),(e) or are separated at a distance $D \cong R_g$, (b),(d),(f). Counterions are gray, neutral monomers are shown in pale colors, and charged monomers in dark colors.

and can facilitate some of the terminal monomers on back-folding. In fact in the case of only terminally charged dendrimers the internal configurations of dendrimers become approximately uniformly charged.[53] The back folding of some of the branches is to a lesser extent also found in the case of neutral dendrimers, where it is attributed to steric interactions. Especially for higher generation dendrimers this effect is often encountered due to the exponential growth of the density with generation shell.

It is not surprising that this visible change in structure is also reflected in the effective interactions between charged dendrimers. The effective force, Eq. (86), can be measured by the same method as discussed for the neutral dendrimers, except that now also the free counterions contribute to the total force acting on the center of mass of each individual dendrimer. In Figure 9 the results of these measurements are shown for the three types of charge distribution, as well as for different spacer lengths. The shape of the forces indicate that the underlying potentials are of Gaussian shape, in agreement with other work published on this topic.[51,52,66]

On raising the overall charge of dendrimers from $Z = 0$ to $Z = 62$, we observe a monotonic increase in the effective force. Depending on the spacer length, the maximum that the force attains lies between two and three times that of the force between neutral dendrimers. At the same time its range is enhanced due to the swelling of the molecule originating from the internal electrostatic repulsive interactions. At large separation, the interaction shows the typical Yukawa-like decay found for charged molecules screened by solvent and counterions. If the mutual distance between the dendrimers is decreased the force rises until a distance of about $D \cong 1.2R_g$ is reached. For shorter distances the force

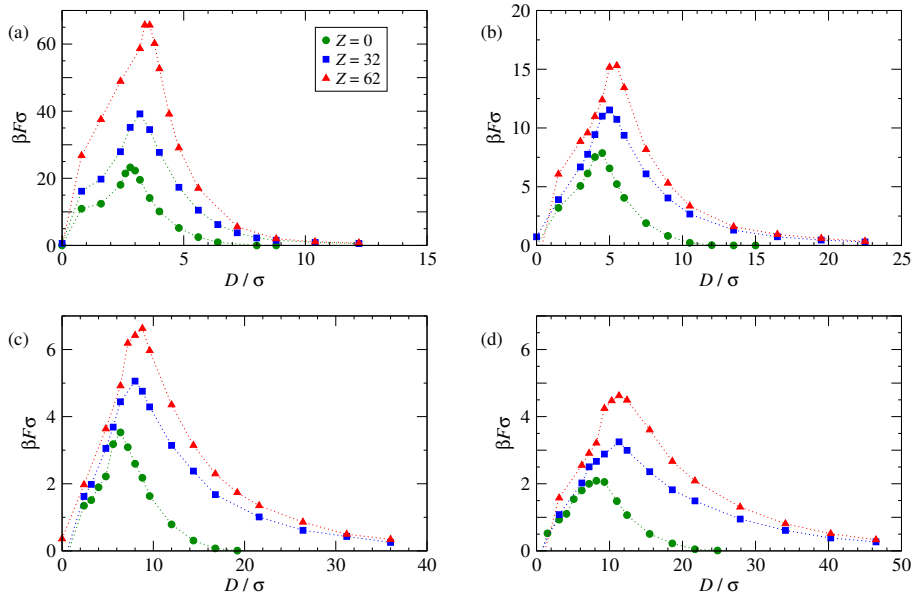


Figure 9: The effective force, F, between $G = 4$ dendrimers with number of spacer segments (a) $P = 1$, (b) $P = 2$, (c) $P = 4$, and (d) $P = 6$ at different pH values of the solvent. The total charge of a dendrimer is denoted by Z.

decreases and drops to zero at zero distance. A change in the spacer length has relatively little effect on these observations. The main consequences of increasing the length of the chain segment between nodes in the dendrimer, is that it results in an increased overall size of the dendrimer and the corresponding change in range of the interaction. Closely related to that, is the fact that this also weakens the strength of the interactions, because the longer spacers allow more freedom to avoid both steric and electrostatic interactions as well as it makes it easier for the counterions to explore the inner regions and screen the bare interactions more efficiently.

5.4 Effect of salt

So far we only discussed the effect of the different charge distributions of dendrimers in an implicit solvent that, in order to maintain an overall charge neutrality, has been completed with the appropriate number of monovalent counterions. We can however still influence the interaction between the macromolecules further by for instance the addition of salt. Within the current model this means the addition of an equal number of positively and negatively charged ions with the same short-range, repulsive Lennard-Jones interaction (81) as all other monomers and counterions. Here we consider three different typical concentrations of the salt corresponding to 0.05, 0.10, and 0.50 $\mathrm{mol/l}$. Note that the number of ion-pairs this corresponds to depends on the size of the simulation box.

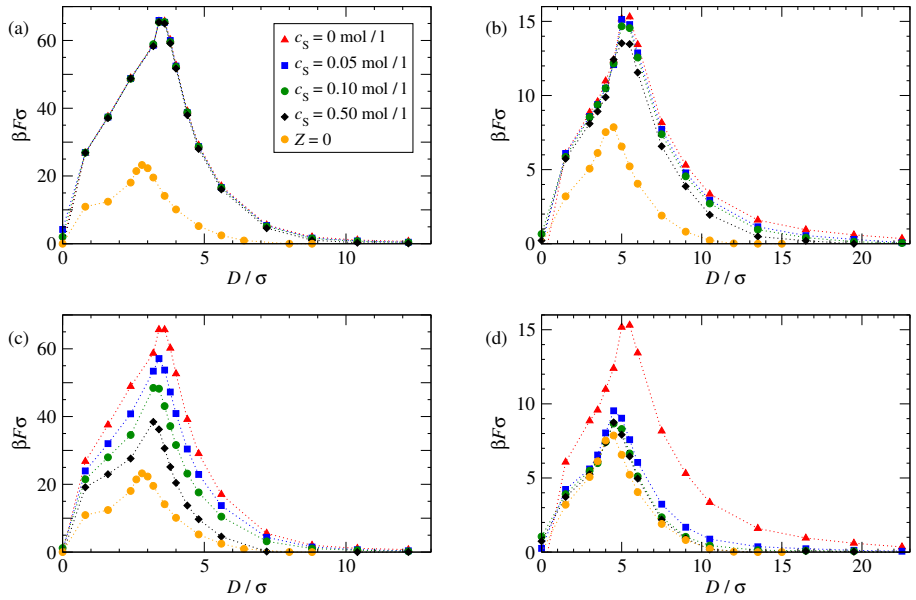


Figure 10: The effective force, F, between fully charged dendrimers with number of spacer segments $P = 1$ [(a),(c)] and $P = 2$ [(b),(d)] at different salt concentrations c_s, where the salt ions are either monovalently charged [(a),(b)] or divalently charged [(c),(d)]. The forces for neutral dendrimers without salt are shown for comparison.

The effective force measured for the different salt concentrations are shown in

Figure 10(a) and (b) for the case of fully charged dendrimers with spacer length $P = 1$ and $P = 2$ respectively. In order to interpret the results, the interaction for uncharged dendrimers are also shown. For the case $P = 1$, we find that there is no dependence on the salt concentration, in fact the four salt concentrations, that is including no added salt, coincide over the whole range. For the longer spacer $P = 2$, we see that for the short distances the force is still unaffected, but for larger distanced decreases slightly upon increasing the salinity. This can be understood in the light of the fact the $P = 1$ dendrimers have a very compact structure that does not allow for additional salt ion-pairs the explore the interior of the structure. By increasing the spacer to $P = 2$, more volume is created that will be accessible to salt molecules, and more so in outer shells of the dendrimer. In so doing, the combined counterions and salt ion-pairs are capable of screening the electrostatic interactions more efficiently and reducing the effective force at larger separations.

A more interesting situation arises when we replace the monovalent salt used above by divalent salt. Upon increasing the concentrations, the effective forces are significantly reduced in the case of spacer length $P = 1$ approaching the interaction of neutral dendrimers. In the case $P = 2$ the forces for the three non-zero concentrations nearly collapse on the neutral case. The divalent ions are much more efficient in screening the interactions than the monovalent ones. In addition, a single divalent ion can replace two monovalent ions. What is happening here is that the monovalent counterions originally found within the dendrimer are expelled and replaced by divalent salt ions, because this reduces the steric interaction in the interior. This is advantageous for another reason, because by having two counterions in bulk rather than a single divalent ion the entropy is increased. The compact structure of the $P = 1$ dendrimer still obstructs this process, but already for the $P = 2$ spacer length this more efficient screening almost eliminates the Coulomb contribution to the effective interaction.

6 Ring Polymers

Ring polymers are macromolecules obtained by joining together the two free ends of a linear polymer chain. They are the most characteristic prototype of topologically constrained molecules,[82] that allow to appreciate how the mere operation of "closing" a linear polymer chain has profound impact on the structural and dynamical properties of single molecules and concentrated solutions of the same alike.

Interest in ring polymers dates many decades ago, as is witnessed, e.g., in three pioneering papers on the subject: in the work of Frank-Kamenetskii *et al.*,[83] the notion of the *topological interaction* between two rings has been introduced and analyzed quantitatively, which arises from the non-concatenation condition of the same; in the work of Grosberg *et al.*,[84] the *crumpled globule model* of rings in the melt has been put forward, making a strong distinction between the structures of linear- and ring-polymer melts; and finally, Obukhov *et al.*[85] have put forward an *annealed lattice-animal* picture of a ring polymer in a melt, deriving thereby novel scaling laws for the diffusion coefficient and the longest relaxation time of a ring, and thereby revising earlier predictions.[86,87] Despite their conceptual simplicity and their highly interesting characteristics, the study of ring polymers, both experimentally and theoretically, is confronted with many obstacles.[88] From the theoretical point of view, the main difficulty of rings in comparison to their linear counterparts lies indeed in the treatment of the topological constraints, which, *inter alia*,

prevents the formulation of the problem in terms of a field-theoretical approach[89] that has proven extremely fruitful for the treatment of solutions or melts of linear chains. On the experimental side, the main problems are related to the difficulty of controlling the synthetic processes so as to obatin monodisperse rings of identical knotedness. Accordingly, computer simulations have emerged as an indispensable tool for the analysis of static and dynamic properties of ring polymers and their solutions.[88,90–113,89] A great deal of interest in ring polymers is motivated by their biological relevance. Indeed, biopolymers such as DNA or chromosomes are often found in a topologically constrained state when they are packed within cells or eukariotes.[114–116,84] In 1993, Grosberg *et al.*[84] demonstrated that molecules that have a topological constraint appear to be able to survive longer in an out-of-equilibrium state that allows for more compact structures, therefore hypothesizing that the long-lasting problem of packing of, e.g., chromosomes in eukaryotes that could not be explained by packing of linear biopolymers, could be solved by adding a topological constraint. Proteins and DNA display a rich variety of topological effects, both structural and dynamic. For example, the DNA of bacteria is present in the traditional double-helix form, but contrarily to what happens for eukaryotes, they have circular chromosomes contained in a DNA helix is closed into a ring.[117] Formation of knots along the backbone of DNA and their location in dependence of the varying rigidity along the backbone of the macromolecule in the bulk[118] and in confinement[119] are another manifestation of the importance of topological concepts for biologically relevant processes.

At the single-molecule level (equivalent to the infinite-dilution limit of a polymer solution), topology manifests itself in various ways. Although the infinite-dilution gyration radius of the rings, $R_{g,0}$, scales with monomer number N with the same, Flory exponent $\nu = 0.588$ as the linear chains in athermal solvents ($R_{g,0} \sim N^{\nu}$), topology effectively expresses itself as a larger excluded-volume parameter, resulting into a lowering of the Θ-temperature of the rings in comparison to that of the linear polymers.[120,111] A related, remarkable effect is the fact that in contrast to *ideal* (i.e., without excluded volume) linear polymers, ideal ring polymers experience an effective repulsion between molecules that is purely due to the additional topological constraint of closing each chain into a loop,[121,112] leading to a scaling $R_{g,0} \sim N^{\nu}$ that is identical to that of self-avoiding rings.

The effects of topology become even stronger at higher concentrations, and in particular at those exceeding the overlap density of the rings. Whereas the concentration screens out the excluded-volume interaction for linear chains, resulting into Gaussian statistics between the correlation blobs of the same,[122] the topological potential between different rings cannot be screened out.

Solutions of ring polymers present a melt viscosity that is lower by one order of magnitude with respect to a solution of linear chains in the same density and solvent conditions.[123–126] Investigations on melts of unknotted, non-concatenated rings[88,93–95,91] have shown that they display a higher diffusivity[88,93–95,91,127,104] and that the Rouse regime extends to larger scales than in their linear counterparts.[100] Rheological experiments[128] and simulations[104] have revealed a power-law stress relaxation, instead of the usual reptation-like exponential behavior found for linear chains. Semiflexible rings, on the other hand, feature a particular form of self-organization in semidilute solutions, forming a disordered state of columnar clusters penetrated by other rings,[112] and displaying an unusual dynamic scenario in which the coherent and the incoherent correlation functions are decoupled from one another, resulting into a state that has been termed *cluster glass.*[113]

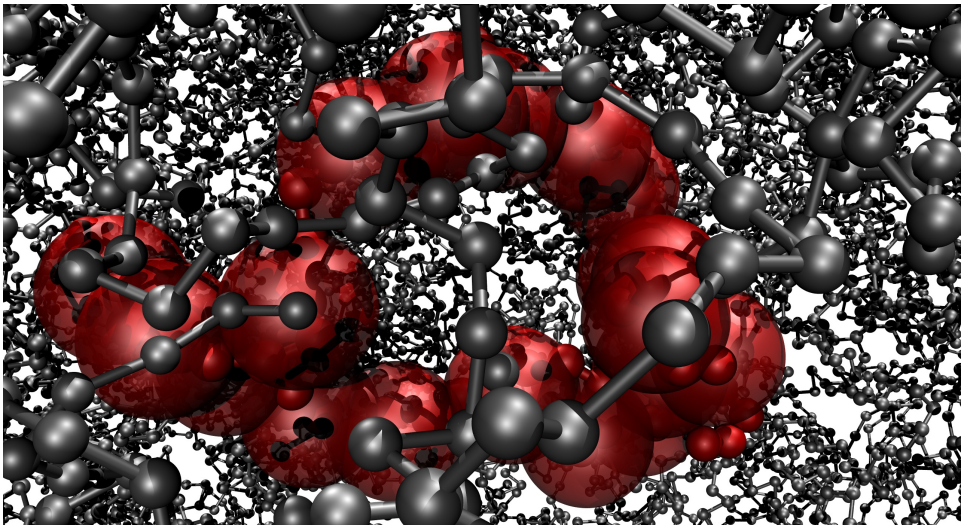


Figure 11: A sketch of a coarse-grained polymer ring chain (red) in a solution of full monomer molecules (grey).

It is evident, thus, that the properties of topologically constrained molecules in the semi-dilute regime are extremely difficult to access, both via theoretical approaches and with computational studies. The difficulty of the latter increases with polymer size and density of polymers in solution, as the bigger the ring and the more rings in solution, the more the monomers to simulate, and the topological tests required in order to preserve the original topology of the system. It therefore becomes of crucial importance being able to analyze, simulate and access the semi-dilute regime for molecules with a ring architecture. A full-monomer detailed representation of semi-dilute solutions, due to both the high number of monomers that it would be necessary to simulate and the high number of topological checks that would be needed, appears to be quite prohibitive if simulations have to be performed for very large polymers in density regimes close and above the overlap concentration, and therefore a coarse-grained approach is called for.

The simplest coarse-graining strategy amounts to replacing the entire ring with *a single effective coordinate*, usually chosen to be the molecule's center of mass,[101,102] in close analogy with the case of linear chains.[34] This single-blob representation of the rings already expresses some of the distinct features related with the ring topology: the resulting effective interaction has a very different amplitude and shape with respect to the Gaussian effective interaction of linear chains,[102] a feature that has been recently rationalized in terms of the strong asymmetry in the sizes of two interpenetrating ring polymers, stemming from the threading of one through the other.[129] Further, the single-blob approach is sensitive to the type of knotedness of the rings, at least for moderate sizes of the same.[102,130,129]

6.1 The full-monomer models

To represent the ring polymers, monomers are modeled as hard-spheres of diameter d and the connections among them are implemented as threads of maximal surface-to-surface

extension δd ($\delta > 1$). Accordingly, the monomer-monomer interaction $V_{\rm mm}(r)$ and the bonding interaction $V_{\rm bond}(r)$, where r is the distance between the monomer centers, read as:

$$V_{\rm mm}(r) = \begin{cases} \infty \text{ for } \frac{r}{d} < 1 \\ 0 \text{ for } \frac{r}{d} > 1 \end{cases} \tag{87}$$

acting among all monomers and

$$V_{\rm bond}(r) = \begin{cases} \infty \text{ for } & \frac{r}{d} < 1 \\ 0 \text{ for } 1 < & \frac{r}{d} < 1+\delta \\ \infty \text{ for } & \frac{r}{d} > 1+\delta \end{cases} \tag{88}$$

for connected ones. Within this representation, we prevent crossing of the bonds of the rings and thus conserve all of the intra- and inter-molecular topology avoiding any accidental concatenations, by setting $\delta = 0.2$ and choosing the Monte Carlo (MC) displacement step to be less or equal to δ.

6.2 Effective pair interactions between two rings

As mentioned in the introduction, properties of ring polymer solutions in the semi-dilute regime are very interesting. If big molecules have to be simulated, both the high number of monomeric units and the topological tests required to simulate the system grow with increasing molecule size and number of molecules in solution till becoming prohibitive. A possible way to simulate those systems is to extract effective interactions between the centre of masses of the molecules and then use such effective interactions to represent the polymer chains. We will in this section deeply analyse how to compute the enter-of-mass effective interactions between two ring polymers with the Widom insertion method[101,102] $V_{\rm eff}(R)$. non interacting two ring polymers, $\mathcal{H}$ have the general form:

$$\mathcal{H} = \mathcal{H}_{11}(\{\mathbf{r}^N\}) + \mathcal{H}_{22}(\{\mathbf{s}^N\}) + \mathcal{H}_{12}(\{\mathbf{r}^N, \mathbf{s}^N\}), \tag{89}$$

where $\{\mathbf{r}^N\}$ and $\{\mathbf{s}^N\}$ are the collective coordinates of the segments of polymer 1 and 2, respectively, $\mathcal{H}_{11}(\{\mathbf{r}^N\})$ and $\mathcal{H}_{22}(\{\mathbf{s}^N\})$ are the intra-ring parts of the Hamiltonian and $\mathcal{H}_{12}(\{\mathbf{r}^N, \mathbf{s}^N\})$ contains all the inter-ring interactions between the segments. The canonical and topologically faithful partition function $\mathcal{Z}^T$ is given as

$$\mathcal{Z}^T = \int\int \mathrm{d}\mathbf{r}^N \mathrm{d}\mathbf{s}^N \exp[-\beta\mathcal{H}]^T. \tag{90}$$

Note the usage of the notation $\mathcal{Z}^T$ and $\exp[-\beta\mathcal{H}]^T$ with the T-superscript. This indicates that, in addition to the usual Boltzmann weight, expressed in the exponential factor of the Hamiltonian, there is an explicit *exclusion* from the partition sum of all microstates that lead to concatenated rings. In what follows, a T-superscript will always be used to indicate the presence of this topological constraint, whereas its absence will denote a usual partition sum, in which concatenated configurations are allowed. Additional constraints on the coordinates can be formally handled by, e.g., introducing appropriate δ-functions in the integrand of Eq. (90). The separation between the centers of mass, for instance, can be

formally fixed at a distance $R = |\mathbf{R}|$ to define the constrained partition function $\mathcal{Z}^T(R)$ as:

$$\mathcal{Z}^T(R) = \int\int \mathrm{d}\mathbf{r}^N \mathrm{d}\mathbf{s}^N \exp[-\beta\mathcal{H}]^T \delta\left(\left|\frac{1}{N}\sum_{i=1}^{N}(\mathbf{r}_i - \mathbf{s}_i)\right| - R\right). \tag{91}$$

With the help of $\mathcal{Z}^T(R)$ given in Eq. (91), the effective interaction $V_{\mathrm{eff}}(R)$ between the centers of mass of the molecules is defined as:

$$\exp\left[-\beta V_{\mathrm{eff}}(R)\right] \equiv \frac{\mathcal{Z}^T(R)}{\mathcal{Z}^T(R\to\infty)} = \frac{\mathcal{Z}^T(R)}{\mathcal{Z}(R\to\infty)}, \tag{92}$$

where we have taken into account that when two finite rings are infinitely far apart from one another, they are also non-concatenated, thus one can set $\mathcal{Z}^T(R\to\infty) = \mathcal{Z}(R\to\infty)$. By multiplying and dividing the right-hand side with $\mathcal{Z}(R)$, one can decompose the effective interaction $V_{\mathrm{eff}}(R)$ into its steric part $V_{\mathrm{steric}}(R)$ and the topological potential $V_T(R)$:

$$V_{\mathrm{eff}}(R) = V_{\mathrm{steric}}(R) + V_T(R), \tag{93}$$

the two terms on the right-hand side being expressed as

$$\exp\left[-\beta V_{\mathrm{steric}}(R)\right] = \frac{\mathcal{Z}(R)}{\mathcal{Z}(R\to\infty)} \tag{94}$$

and

$$\exp\left[-\beta V_T(R)\right] = \frac{\mathcal{Z}^T(R)}{\mathcal{Z}(R)}. \tag{95}$$

The steric potential $V_{\mathrm{steric}}(R)$, Eq. (94), is thus expressed in the usual way in which effective interactions between macromolecular entities are defined in cases where topology plays no role. Eq. (93) above demonstrates however, that for ring polymers, $V_{\mathrm{steric}}(R)$ is only part of the full story: an additional, topological term $V_T(R)$ must be added to it to obtain the full effective potential $V_{\mathrm{eff}}(R)$. Though, as it will be shortly shown, for two rings $V_T(R)$ is only a small fraction of $V_{\mathrm{eff}}(R)$, in concentrated solutions where steric interactions are increasingly screened out, the topological interaction plays a very important role in determining the conformations of the molecules and the correlations between them.

To compute the two contributions to the effective pair potential in the simulation, we use a generalization of the Widom insertion algorithm.[34,22,35] Our choice is guided by the fact that as $R\to\infty$, and since the inter-monomer potentials are short-ranged, the interaction term $\mathcal{H}_{12}(\{\mathbf{r}^N, \mathbf{s}^N\})$ vanishes identically and the denominator of Eq. (92) factorizes into the product of the partition functions of two noninteracting rings. Concomitantly, Eq. (92) can be re-expressed as

$$\exp\left[-\beta V_{\mathrm{eff}}(R)\right] = \tag{96}$$
$$\left\langle \exp\left[-\beta\mathcal{H}_{12}(\{\mathbf{r}^N, \mathbf{s}^N\})\right]^T \delta\left(\left|\tfrac{1}{N}\textstyle\sum_{i=1}^{N}(\mathbf{r}_i - \mathbf{s}_i)\right| - R\right)\right\rangle_{\mathrm{intra}},$$

where the notation $\langle\cdots\rangle_{\mathrm{intra}}$ indicates that the expectation value has to be calculated in the ensemble of the non-interacting intramolecular Hamiltonian $\mathcal{H}_{\mathrm{intra}}$ of two independent rings:

$$\mathcal{H}_{\mathrm{intra}} = \mathcal{H}_{11}(\{\mathbf{r}^N\}) + \mathcal{H}_{22}(\{\mathbf{s}^N\}). \tag{97}$$

It might appear at first sight paradoxical that one is capable of expressing a constrained free energy as the expectation value of some quantity. However, as Eq. (92) readily shows, $V_{\rm eff}(R)$ is a *difference* between two constrained free energies, one at separation R and the other at infinite separation. Free energy differences can indeed be calculated very efficiently in computer simulations.

As mentioned in 3.3 the Widom insertion method has proven successfull in extracting effective itneractions between macromolecules. In the case of ring polymers, it must be emphasized that the arbitrary insertion of one ring in the neighborhood of another always entails the risk of producing a concatenated pair. Therefore, due care has to be taken *a posteriori* to exclude such cases from the calculation of the expectation value in Eq. (97).

For the case at hand, we proceeded as follows. For each insertion step we randomly select two molecules from an ensemble of $L = 10^4$ isolated ring equilibrium configurations. We then compute, for all possible $\mathcal{L} = L(L-1)/2$ combinations of the equilibrium configurations, the probability $\mathcal{P}_{TS}(R)$ of bringing the centers of mass of the two molecules at a distance R, under the condition that both the topological (T) and the steric (S) constraints are fulfilled. To this end, the following procedure has been followed. Each microstate $\mu(R)$ in which the centers of mass of the two rings are separated by R is first checked for steric interactions and it is provisionally accepted with a steric hindrance acceptance probability $p_S^{\rm acc}(\mu(R))$ given by:

$$p_S^{\rm acc}(\mu(R)) = \exp[-\beta \mathcal{H}_{12}(\mu(R))], \tag{98}$$

which never exceeds unity, since all cross-ring interactions are repulsive. In this way, $\mathcal{L}_S(R)$ out of $\mathcal{L}$ configurations survive the steric test.[d]

To account for inter-winding of the two rings, we compute the Gauss linking number, m, which is a measure of the degree of concatenation of two molecules.[121,101] For any microstate $\mu(R)$, m is given by

$$m(\mu(R)) = \frac{1}{4\pi} \oint_{C_1} \oint_{C_2} \frac{(\mathrm{d}\mathbf{r} \times \mathrm{d}\mathbf{s}) \cdot (\mathbf{r} - \mathbf{s})}{|\mathbf{r} - \mathbf{s}|^3}, \tag{99}$$

where the $\mathbf{r}$- and $\mathbf{s}$-integrations run along the closed contours of the ring polymers C_1 and C_2, respectively. When $m \neq 0$ the two molecules are concatenated, while $m = 0$ implies non-concatenation.[e] Accordingly, the probability of acceptance for the combination of steric *and* topological constraints, $p_{TS}^{\rm acc}(\mu(R))$, reads as

$$p_{TS}^{\rm acc}(\mu(R)) = \delta_{m,0} p_S^{\rm acc}(\mu(R)). \tag{100}$$

In this way, the number of accepted configurations is further reduced from $\mathcal{L}_S(R)$ to $\mathcal{L}_{TS}(R)$ and the effective interaction is computed as

$$\beta V_{\rm eff}(R) = -\ln\left[\frac{\mathcal{L}_{TS}(R)}{\mathcal{L}}\right] = -\ln\left[\mathcal{P}_{TS}(R)\right]. \tag{101}$$

[d] Note that for Model I the exponential factor at the right-hand side of Eq. (98) is either zero, if any two monomers of the two rings overlap, or unity, otherwise, and thus the survival of a microstate is not any more a matter of chance.

[e] This is not entirely correct. Although $m \neq 0$ always implies the existence of concatenation, the opposite is not true, since there are some particular situations, such as the Whitehead link, for which $m = 0$. We have ignored these special cases here.

The steric part, $V_{\rm steric}(R)$, is calculated as $-k_{\rm B}T$ times the logarithm of the probability $\mathcal{P}_S(R)$ of passing the steric requirements, i.e.,

$$\beta V_{\rm steric}(R)] = -\ln\left[\frac{\mathcal{L}_S(R)}{\mathcal{L}}\right] = -\ln\left[\mathcal{P}_S(R)\right]. \tag{102}$$

The topological potential is now easily obtained with the help of Eqs. (93), (101) and (102) as $-k_{\rm B}T$ times the logarithm of the ratio of the number of configurations fulfilling the combination of T- and S-conditions over those fulfilling only S:

$$\beta V_T(R) = -\ln\left[\frac{\mathcal{L}_{TS}(R)}{\mathcal{L}_S(R)}\right] = -\ln\left[\mathcal{P}_{T|S}(R)\right]. \tag{103}$$

In the last equation, we introduced the conditional probability $\mathcal{P}_{T|S}(R)$ that a given configuration will fulfill the topological constraints *provided* it fulfills the steric ones. This follows immediately from Eqs. (101)-(103) above as a corollary of the rule for calculating conditional probabilities:

$$\mathcal{P}_{T|S}(R) = \frac{\mathcal{P}_{TS}(R)}{\mathcal{P}_S(R)}. \tag{104}$$

For purely repulsive monomers, the contribution $V_{\rm steric}(R)$ is positive, since the proximity of the two molecules restricts the number of conformers for both, thereby reducing the entropy of the system. This is also evident from Eq. (102) above, valid *only* for repulsive interactions, which expresses $\beta V_{\rm steric}(R)$ as minus the logarithm of a probability and it thus implies $\beta V_{\rm steric}(R) \geq 0$ for all R. However, in solvents of worsening quality, for which enthalpic terms would be present in the intermonomer interactions, this quantity can indeed develop attractive parts for some ranges of the intermolecular separation.[22,111] The topological potential, on the other hand, is given as $\beta V_T(R) = -\ln\left[\mathcal{Z}^T(R)/\mathcal{Z}(R)\right]$, and the numerator of the ratio within the logarithm is *always* smaller than its denominator, since those microstates that violate the non-concatenation condition are included in the partition sum $\mathcal{Z}(R)$ but excluded from $\mathcal{Z}^T(R)$. Topology sets forth a constraint that inadvertently reduces the number of permissible microstates, therefore the topologcal contribution $\beta V_T(R)$ is non-negative at all R, independently of the microscopic details of the model. Still, a weakly attractive, topological effective force $\mathbf{F}_T(R) = -\nabla V_T(R)$ can emerge between two rings at moderately small values of their separation, $R \leq R_{g,0}/2$, see below.

Results for the effective and topological potentials $V_{\rm eff}(R)$ and $V_T(R)$ are reported in Figure 12 The topological effective potentials obtainedl are quantitatively comparable to those predicted by Hirayama *et al.*,[121] who employed the self-avoiding polygon model, by Bohn and Heermann,[101] who performed lattice simulations, as well as by Narros *et al.*[102] in their off-lattice simulations. As predicted by the various authors, $V_T(R)$ is positive and it displays a maximum at $R \cong R_{g,0}/2$, which accounts for about 13% of the maximum value of the total effective potential.

Coarse-graining a ring polymer with a single effective potential has been proven to be a good approach up to densities of the order of the overlap density ρ^*. As soon as the concentration approaches ρ^*, the single blob representation looses its validity and many body effects start playing an important role.[102] It is possible to overcome the shortcomings of the single-blob representatiob by switching to a so called "multi-blob" representation:[5,36]

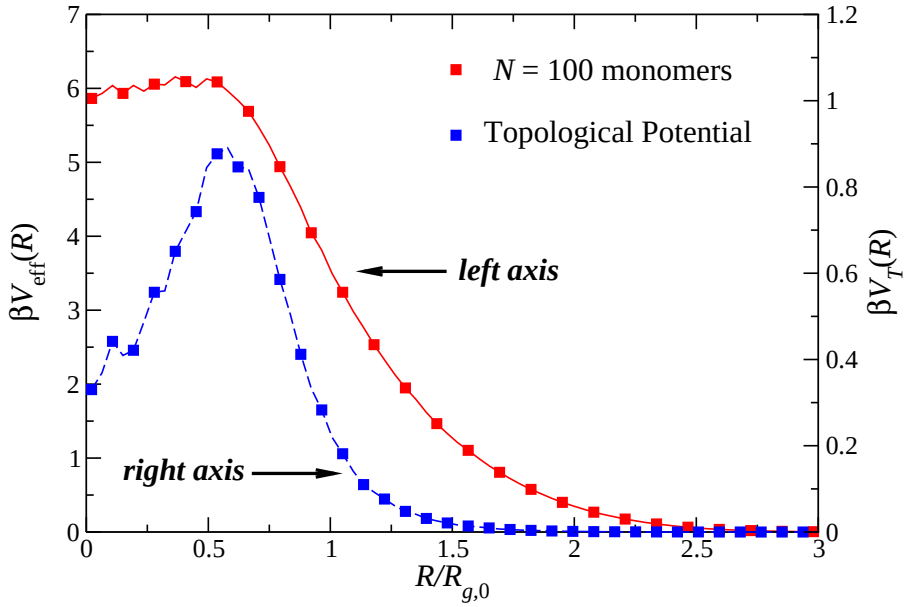


Figure 12: The effective potential $V_{\text{eff}}(R)$ (left vertical axis) and the topological potential $V_T(R)$ (right vertical axis) between the centers of mass of two flexible, unknotted and non-concatenated ring polymers.

each ring made of N monomers will be "broken into" a number of n_B blobs, every blob containing $g = N/n_B$ monomers. The number of blobs is chosen in such a way that each coarse-grained entity is sufficiently small to be below its own overlap concentration, so that many body effects play a minimum role and only pair overlaps are significant. Such a coarse-graining model is extremely powerful and allows to explore the semi-dilute regime of ring polymers without any loss of information with respect to full monomer simulations. Static properties, such as the pair distribution funciton can be obtained with this coarse-graining methodology for various densities and, as shown in Figure 13, the agreement between the two representations is striking. It is important to stress that the multiblob methodology allows to simulate arbitrarily long polymers at a small fraction of the computational cost that would have been required in a full-monomer representation. For more details, we refer the reader to Ref.[36] and references therein.

7 Concluding Remarks

We have presented a concise exposure to the strategy of coarse-graining of complex fluids by means of systematic elimination ("integrating out") of degrees of freedom that are small and fast in comparison to the ones in which we are interested. It must be emphasized, of

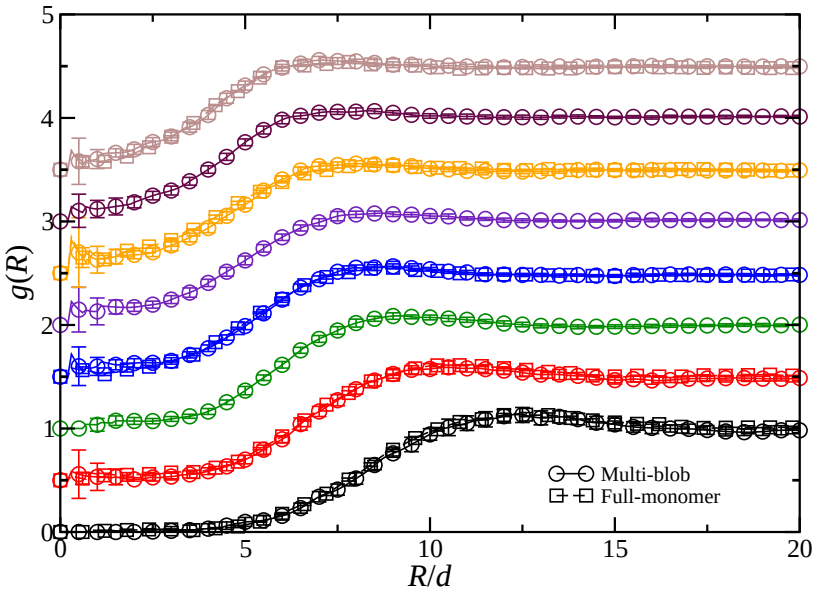


Figure 13: The pair distribution functions $g(R)$ between the centers of mass of unknotted, flexible and non-concatenated ring polymers at various polymer densities ρ/ρ_*. Shown are results from both the full-monomer simulations of Model I with $N = 100$ hard monomers per ring (for $\rho/\rho_* = 0.5, 1.0, 2.0, 3.0$ and 4.0) and the multi-blob simulation with $n_B = 50$ blobs per ring. From bottom to top, densities are increased in steps of $\Delta\rho = 0.5\rho_*$, starting from the value $\rho = 0.5\rho_*$. For clarity, each curve has been shifted up by an amount of 0.5 from the preceding one.

course, that their presence has a profound influence in the structure and the dynamics of the bigger and slower degrees of freedom: therefore, they might be uninteresting but they are not irrelevant, and therefore they cannot be ignored! The general formalism has been put forward, which is valid for any classical system in thermodynamic equilibrium. In fact, the formalism leading to the derivation of the effective Hamiltonian does not distinguish at all between the small/fast and the big/slow degrees of freedom: the very same framework is equally valid in case one is interested in eliminating the big degrees of freedom in favor of the small. This unusual state of affairs is, however, just an academic exercise: on the one hand, it is of little or no practical advantage and on the other the effective interactions that would result in such a case would be heavily many-body in nature.

We have also presented some concrete examples of coarse-graining, in which a full bridging of the length scales has been achieved: from the monomer-monomer interactions (microscopic) to the effective interactions and structure (mesoscopic) all the way to the thermodynamic phase behavior (macroscopic). The microscopic interactions set forth were of some generic nature, based on a non-specific modeling of the steric and the dispersion

forces. A more detailed calculation there would begin at the level of quantum mechanics, and the solution of a many-body electronic-structure problem would be required to derive effective atom-atom potentials starting, essentially, from the Schrödinger equation,[131] a topic which is beyond the scope of this Chapter. Finally, we emphasize that the contents of this Chapter are limited exclusively to equilibrium properties. For extensions of the notion of the effective interaction to systems out of equilibrium, we refer the reader to the literature[132] and references therein.

Acknowledgments

B. C. acknowledges support from the Austrian Academy of Sciences (ÖAW) through her APART Fellowship No. 11723. L. R. acknowledges support from the Austrian Research Fund (FWF) through his Lise-Meitner Fellowship M No. 1650-N27. We would like to thank a number of coworkers for exciting and productive collaborations on the topics related to this Chapter, namely: Manuel Camargo, Sebastian Huißmann, Christian Koch, Federica Lo Verso, Daniela Marzi, Angel Moreno, Arturo Narros, Nicoletta Gnan, Athanassios Z. Panagiotopoulos, Domenico Truzzolillo, Dimitris Vlassopoulos and Emanuela Zaccarelli. Fruitful and numerous discussions with Emanuela Bianchi, Thiago Escobar Colla, Daan Frenkel, Gerhard Kahl, Emanuele Locatelli, Bianca Mladek and Peter Poier are also gratefully acknowledged.

References

1. C. N. Likos, *Effective interactions in soft condensed matter physics*, Phys. Rep., **348**, 267, 2001.
2. P.-G. De Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, N.Y., 1979.
3. P. J. Flory and W. R. Krigbaum, *Statistical mechanics of dilute polymer solutions*, J. Chem. Phys, **18**, 1086, 1950.
4. A. V. Grosberg, P. G. Khalatur, and A. R. Khoklov, *Polymeric coils with excluded volume in dilute solution: The invalidity of the model of impenetrable spheres and the influence of excluded volume on the rates of diffusion-controlled intermacromolecular reactions*, Makromol. Chem. Rapid. Commun., **3**, 709, 1982.
5. C. Pierleoni, B. Capone, and J.-P. Hansen, *A soft effective segment representation of semidilute polymer solutions*, J. Chem. Phys., **127**, 171102, 2007.
6. A. A. Louis, P. G. Bolhuis, R. Finken, V. Krakoviack, E. J. Meijer, and J. P. Hansen, *Coarse-graining polymers as soft colloids*, Physica A, **306**, 251 – 261, 2002.
7. E. Bianchi, R. Blaak, and C. N. Likos, *Patchy colloids: state of the art and perspectives*, Phys. Chem. Chem. Phys., **13**, 6397–6410, 2011.
8. J. P. K. Doye, T. E. Ouldridge, A. A. Louis, F. Romano, P. Sulc, C. Matek, B. E. K. Snodin, L. Rovigatti, J. S. Schreck, R. M. Harrison, and W. P. J. Smith, *Coarse-graining DNA for simulations of DNA nanotechnology*, Phys. Chem. Chem. Phys., **15**, 20395–20414, 2013.
9. J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman, *The Theory of Critical Phenomena: an Introduction to the Renormalization Group*, Oxford University Press, Inc., New York, 1992.

10. M. Dijkstra, R. van Roij, and R. Evans, *Phase diagram of highly asymmetric binary hard-sphere mixtures*, Phys. Rev. E, **59**, 5744–5771, 1999.
11. B. Widom, *Structure of interfaces from uniformity of the chemical potential*, J. Stat. Phys., **19**, 563–574, 1978.
12. S. Asakura and F. Oosawa, *How properties of interacting depletant particles control aggregation of hard-sphere colloids*, J. Chem. Phys., **22**, 1255, 1954.
13. S. Asakura and F. Oosawa, *Interaction between particles suspended in solutions of macromolecules*, J. Polym. Sci., **33**, 183–192, 1958.
14. A. Vrij, *Polymers at interfaces and interactions in colloidal dispersions*, Pure Appl. Chem., **48**, 471, 1976.
15. Y. Mao, M. E. Cates, and H. N. W. Lekkerkerker, *Depletion force in colloidal systems*, Physica A, **222**, 10–24, 1995.
16. K. Binder, P. Virnau, and A. Statt, *Perspective: The Asakura-Oosawa model: A colloid prototype for bulk and interfacial phase behavior*, J. Chem. Phys., **141**, 140901, 2014.
17. J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, Elsevier, Amsterdam, 2nd edition, 1990.
18. W. B. Russel, D. A. Saville, and W. R. Schowalter, *Colloidal Dispersions*, Cambridge University Press, Cambridge, 1992.
19. L. Belloni, *Colloidal interactions*, J. Phys.: Condens. Matter, **12**, R549, 2000.
20. C. Von Ferber, A. Jusufi, C. N. Likos, H. Löwen, and M. Watzlawek, *Triplet interactions in star polymer solutions*, Eur. Phys. J. E, **2**, 311–318, 2000.
21. P. G. Bolhuis, A. A. Louis, and J. P. Hansen, *Many-body interactions and correlations in coarse-grained descriptions of polymer solutions*, Phys. Rev. E, **64**, 021801, 2001.
22. V. Krakoviack, J.-P. Hansen, and A. A. Louis, *Influence of solvent quality on effective pair potentials between polymers in solution*, Phys. Rev. E, **67**, 041801, 2003.
23. G. M. Torrie and J. P. Valleau, *Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling*, J. Comp. Phys., **23**, 187–199, 1977.
24. D. Marzi, C. N. Likos, and B. Capone, *Coarse graining of star-polymer – colloid nanocomposites*, J. Chem. Phys., **137**, 014902, 2012.
25. F. Wang and D. P. Landau, *Efficient, multiple-range random walk algorithm to calculate the density of states*, Phys. Rev. Lett., **86**, 2050, 2001.
26. B. M. Mladek, G. Kahl, and C. N. Likos, *Computer assembly of cluster-forming amphiphilic dendrimers*, Phys. Rev. Lett., **100**, 028301, 2008.
27. L. Rovigatti, N. Gnan, A. Parola, and E. Zaccarelli, *How soft repulsion enhances the depletion mechanism*, Soft Matter, **11**, 692–700, 2015.
28. A. M. Ferrenberg and R. H. Swendsen, *New Monte Carlo technique for studying phase transitions*, Phys. Rev. Lett., **61**, 2635, 1988.
29. S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, *The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method*, J. Comp. Chem., **13**, 1011–1021, 1992.
30. J. Dzubiella, H. Löwen, and C. N. Likos, *Depletion forces in nonequilibrium*, Phys. Rev. Lett., **91**, 248301, 2003.
31. J. Dzubiella, C. N. Likos, and H. Löwen, *Phase behavior and structure of star-polymer–colloid mixtures*, J. Chem. Phys., **116**, 9518–9530, 2002.
32. S. Huißmann, C. N. Likos, and R. Blaak, *Effective interactions between charged dendrimers*, Soft Matter, **7**, 8419–8427, 2011.

33. D. A. Lenz, R. Blaak, and C. N. Likos, *Structural properties of dendrimer–colloid mixtures*, J. Phys.: Condens. Matter, **24**, 284119, 2012.
34. P. G. Bolhuis, A. A. Louis, J. P. Hansen, and E. J. Meijer, *Accurate effective pair potentials for polymer solutions*, J. Chem. Phys., **114**, 4269, 2001.
35. B. M. Mladek and D. Frenkel, *Pair interactions between complex mesoscopic particles from Widom's particle-insertion method*, Soft Matter, **7**, 1450, 2011.
36. A. Narros, C. N. Likos, A. J. Moreno, and B. Capone, *Multi-blob coarse graining for ring polymer solutions*, Soft Matter, **10**, 9601–9614, 2014.
37. W. C. K. Poon, J. S. Selfe, M. B. Robertson, S. M. Ilett, A. D. Pirie, and P. N. Pusey, *An experimental study of a model colloid-polymer mixture*, J. Phys. II (Paris), **3**, 1075–1086, 1993.
38. S. Sukenik, L. Sapir, and D. Harries, *Balance of enthalpy and entropy in depletion forces*, Curr. Opin. Coll. Interface Sci., **18**, 495 – 501, 2013.
39. P. J. Lu, E. Zaccarelli, F. Ciulla, A. B. Schofield, F. Sciortino, and D. A. Weitz, *Gelation of particles with short-range attraction*, Nature, **453**, 499–503, 2008.
40. S. Buzzaccaro, J. Colombo, A. Parola, and R. Piazza, *Critical depletion*, Phys. Rev. Lett., **105**, 198301, 2010.
41. N. Gnan, E. Zaccarelli, and F. Sciortino, *Casimir-like forces at the percolation transition*, Nat. Comm., **5**, 3267, 2014.
42. R. Roth, R. Evans, and A. A. Louis, *Theory of asymmetric nonadditive binary hard-sphere mixtures*, Phys. Rev. E, **64**, 051202, 2001.
43. A. A. Louis, P. G. Bolhuis, J-P. Hansen, and E. J. Meijer, *Can Polymer Coils Be Modeled as Soft Colloids?*, Phys. Rev. Lett., **85**, 2522, 2000.
44. E. Buhleier, W. Wehner, and F. Vögtle, *Cascade- and Nonskid-Chain-like Syntheses of Molecular Cavity Topologies*, Synthesis, **2**, 155, 1978.
45. X. Li, M. Zamponi, K. Hong, C.-Y. Shew L. Porcar, T. Jenkins, E. Liu, G. S. Smith, K. W. Herwig, Y. Liu, and W.-R. Chen, *pH Responsiveness of polyelectrolyte dendrimers: a dynamical perspective*, Soft Matter, **7**, 618–622, 2011.
46. A. D'Emanuele and D. Attwood, *Dendrimer–drug interactions*, Adv. Drug Deliv. Rev., **57**, 2147–2162, 2005.
47. N. G. Portney and M. Ozkan, *Nano-oncology: drug delivery, imaging, and sensing*, Anal. Bioanal. Chem., **384**, 620–630, 2006.
48. U. Gupta, H. B. Agashe, A. Asthana, and N. K. Jain, *Dendrimers: Novel Polymeric Nanoarchitectures for Solubility Enhancement*, Biomacromolecules, **7**, 649–658, 2006.
49. R. W. J. Scott, O. M. Wilson, and R. M. Crooks, *Synthesis, Characterization, and Applications of Dendrimer-Encapsulated Nanoparticles*, Journal of Physical Chemistry B, **109**, 692–704, 2005.
50. M. Ballauff and C. N. Likos, *Dendrimers in Solution: Insight from Theory and Simulation*, Angew. Chem. Intl. Engl. Ed., **43**, 2998–3020, 2004.
51. C. N. Likos, M. Schmidt, H. Löwen, M. Ballauff, D. Pötschke, and P. Lindner, *Soft interaction between dissolved flexible dendrimers: theory and experiment*, Macromolecules, **34**, 2914–2920, 2001.
52. C. N. Likos, S. Rosenfeldt, N. Dingenouts, M. Ballauff, P. Lindner, N. Werner, and F. Vögtle, *Gaussian effective interaction between flexible dendrimers of fourth generation: a theoretical and experimental study*, J. Chem. Phys., **117**, 1869–1877, 2002.

53. S. Huißmann, C. N. Likos, and R. Blaak, *Conformations of high-generation dendritic polyelectrolytes*, J. Mater. Chem., **20**, 10486–10494, 2010.
54. G. E. Buhleier, W. Wehner, and F. Vögtle, *"Cascade"- and 'Nonskid-Chain-like" Syntheses of Molecular Cavity Topologies*, Synthesis (Stuttgart), **9**, 155–158, 1978.
55. D. A. Tomalia, A. M. Naylor, and W. A. Goddard III, *Starburst Dendrimers: Molecular-Level Control of Size, Shape, Surface Chemistry, Topology, and Flexibility from Atoms to Macroscopic Matter*, Angew. Chem. Intl. Engl. Ed., **29**, 138–175, 1990.
56. P. Antoni, D. Nyström, C. J. Hawker, A. Hult, and M. Malkoch, *A chemoselective approach for the accelerated synthesis of well-defined dendritic architectures*, Chem. Commun., **22**, 2249–2251, 2007.
57. J. Haensler and F. C. Szoka Jr., *Polyamidoamine cascade polymers mediate efficient transfection of cells in culture*, Bioconjugate Chem., **4**, 372–379, 1993.
58. W.-R. Chen, L. Porcar, Y. Liu, P. D. Butler, and L. J. Magid, *Small Angle Neutron Scattering Studies of the Counterion Effects on the Molecular Conformation and Structure of Charged G4 PAMAM Dendrimers in Aqueous Solutions*, Macromolecules, **40**, 5887–5898, 2007.
59. T. K. Li, K. Hong, L. Porcar, R. Verduzco, P. D. Butler, G. S. Smith, Y. Liu, and W.-R. Chen, *Assess the Intramolecular Cavity of a PAMAM Dendrimer in Aqueous Solution by Small-Angle Neutron Scattering*, Macromolecules, **41**, 8916–8920, 2008.
60. L. Porcar, Y. Liu, R. Verduzco, K. Hong, P. D. Butler, L. J. Magid, G. S. Smith, and W.-R. Chen, *Structural Investigation of PAMAM Dendrimers in Aqueous Solutions Using Small-Angle Neutron Scattering: Effect of Generation*, J. Phys. Chem. B, **112**, 14772–14778, 2008.
61. L. Porcar, K. Hong, P. D. Butler, K. W. Herwig, G. S. Smith, Y. Liu, and W.-R. Chen, *Intramolecular Structural Change of PAMAM Dendrimers in Aqueous Solutions Revealed by Small-Angle Neutron Scattering*, J. Phys. Chem. B, **114**, 1751–1756, 2010.
62. H. M. Harreis, C. N. Likos, and M. Ballauff, *Can dendrimers be viewed as compact colloids? A simulation study of the fluctuations in a dendrimer of fourth generation*, J. Chem. Phys., **118**, 1979–1988, 2003.
63. S. Rosenfeldt, N. Dingenouts, M. Ballauff, P. Lindner, C. N. Likos, N. Werner, and F. J. Vögtle, *Determination of the structure factor of polymeric systems in solution by small-angle scattering: a SANS-study of a dendrimer of fourth generation*, Macromol. Chem. Phys., **203**, 1995–2004, 2002.
64. D. Pötschke, M. Ballauff, P. Lindner, M. Fischer, and F. Vögtle, *Analysis of the Structure of Dendrimers in Solution by Small-Angle Neutron Scattering Including Contrast Variation*, Macromolecules, **32**, 4079–4087, 1999.
65. D. Pötschke, M. Ballauff, P. Lindner, M. Fischer, and F. Vögtle, *The structure of dendritic molecules in solution as investigated by small-angle neutron scattering*, Macromol. Chem. Phys., **201**, 330–339, 2000.
66. I. O. Götze, H. M. Harreis, and C. N. Likos, *Tunable effective interactions between dendritic macromolecules*, J. Chem. Phys., **120**, 7761–7771, 2004.
67. I. O. Götze and C. N. Likos, *Microscopic and coarse-grained correlations of concentrated dendrimer solutions*, J. Phys.: Condens. Matter, **17**, S1777–S1797, 2005.
68. I. O. Götze, A. J. Archer, and C. N. Likos, *Structure, phase behaviour and inhomogeneous fluid properties of binary dendrimer mixtures*, J. Chem. Phys., **124**, 084901, 2006.

69. A. N. Rissanou, I. G. Economou, and A. Z. Panagiotopoulos, *Monte Carlo Simulation of the Phase Behavior of Model Dendrimers*, Macromolecules, **39**, 6298–6305, 2006.
70. G. Nisato, R. Ivkov, and E. J. Amis, *Size Invariance of Polyelectrolyte Dendrimers*, Macromolecules, **33**, 4172–4176, 2000.
71. A. A. Gurtovenko, S. V. Lyulin, M. Karttunen, and I. Vattulainen, *Molecular dynamics study of charged dendrimers in salt-free solution: Effect of counterions*, J. Chem. Phys., **124**, 094904, 2006.
72. J. S. Klos and J.-U. Sommer, *Simulations of Terminally Charged Dendrimers with Flexible Spacer Chains and Explicit Counterions*, Macromolecules, **43**, 4418–4427, 2010.
73. R. Blaak, S. Lehmann, and C. N. Likos, *Charge-induced conformational changes of dendrimers*, Macromolecules, **41**, 4452–4458, 2008.
74. S. Huißmann, A. Wynveen, C. N. Likos, and R. Blaak, *The effects of pH, salt and bond stiffness on charged dendrimers*, J. Phys.: Condens. Matter, **22**, 232101, 2010.
75. S. Huißmann, C. N. Likos, and R. Blaak, *Effective interactions between charged dendrimers*, Soft Matter, **7**, 8419–8427, 2011.
76. D. Frenkel and B. Smit, *Understanding Molecular Simulation*, Academic Press, San Diego, 1996.
77. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press Oxford, 1989.
78. J. D. Weeks, D. Chandler, and H. Andersen, *Role of repulsive forces in determining the equilibrium structure of simple liquids*, J. Chem. Phys., **54**, 5237–5247, 1971.
79. G. S. Grest, K. Kremer, and T. A. Witten, *Structure of many arm star polymers: a molecular dynamics simulation*, Macromolecules, **20**, 1376–1383, 1987.
80. G. S. Grest and K. Kremer, *Molecular dynamics simulation for polymers in the presence of a heat bath*, Phys. Rev. A, **33**, 3628–3631, 1986.
81. J. des Cloizeaux and G. Jannink, *Polymers in Solution*, The Clarendon Press, Oxford, 1990.
82. C. Micheletti, D. Marenduzzo, and E. Orlandini, *Polymers with spatial or topological constraints: theoretical and computational results*, Phys. Rep., **504**, 1, 2011.
83. M. D. Frank-Kamenetskii, A. V. Lukashin, and A. V. Vologodskii, *Statistical mechanics and topology of polymer chains*, Nature, **258**, 398, 1975.
84. A. Grosberg, Y. Rabin, S. Havlin, and A. Neer, *Crumpled Globule Model of the Three-Dimensional Structure of DNA*, Europhys. Lett., **23**, 373, 1993.
85. S. P. Obukhov, M. Rubinstein, and T. Duke, *Dynamics of a ring polymer in a gel*, Phys. Rev. Lett., **73**, 1263, 1994.
86. M. E. Cates and J. M. Deutsch, *Conjectures on the statistics of ring polymers*, J. Phys. France, **47**, 2121, 1986.
87. M. Rubinstein, *Dynamics of ring polymers in the presence of fixed obstacles*, Phys. Rev. Lett., **57**, 3023, 1986.
88. M. Müller, J. P. Wittmer, and M. E. Cates, *Topological effects in ring polymers: A computer simulation study*, Phys. Rev. E, **53**, 5063, 1996.
89. A. Y. Grosberg, *Annealed lattice animal model and Flory theory for the melt of non-concatenated rings: towards the physics of crumpling*, Soft Matter, **10**, 560, 2014.
90. T. Pakula and S. Geyler, *Cooperative relaxations in condensed macromolecular systems. 3. Computer-simulated melts of cyclic polymers*, Macromolecules, **21**, 1665, 1988.

91. M. Müller, J. P. Wittmer, and M. E. Cates, *Topological effects in ring polymers. II. Influence of persistence length*, Phys. Rev. E, **61**, 4078, 2000.
92. M. Müller, J. P. Wittmer, and J.-L. Barrat, *On two intrinsic length scales in polymer physics: Topological constraints vs. entanglement length*, Europhys. Lett., **52**, 406, 2000.
93. S. Brown and G. Szamel, *Structure and dynamics of ring polymers*, J. Chem. Phys., **108**, 4705, 1998.
94. S. Brown and G. Szamel, *Computer simulation study of the structure and dynamics of ring polymers*, J. Chem. Phys., **109**, 6184, 1998.
95. S. Brown, T. Lenczycki, and G. Szamel, *Influence of topological constraints on the statics and dynamics of ring polymers*, Phys. Rev. E, **63**, 052801, 2001.
96. K. Hur, R. G. Winkler, and D. Y. Yoon, *Comparison of ring and linear polyethylene from molecular dynamics simulations*, Macromolecules, **39**, 3975, 2006.
97. B. V. S. Iyer, S. Shanbhag, V. A. Juvekar, and A. K. Lele, *Self-diffusion coefficient of ring polymers in semidilute solution*, J. Polym. Sci. Pol. Phys., **46**, 2370, 2008.
98. T. Vettorel, A. Y. Grosberg, and K. Kremer, *Statistics of polymer rings in the melt: a numerical simulation study*, Phys. Biol., **6**, 025013, 2009.
99. J. Suzuki, A. Takano, T. Deguchi, and Y. Matsushita, *Dimension of ring polymers in bulk studied by Monte-Carlo simulation and self-consistent theory*, J. Chem. Phys., **131**, 144902, 2009.
100. G. Tsolou, N. Stratikis, C. Baig, P. S. Stephanou, and V. G. Mavrantzas, *Melt structure and dynamics of unentangled polyethylene rings: Rouse theory, atomistic molecular dynamics simulation, and comparison with the linear analogues*, Macromolecules, **43**, 10692, 2010.
101. M. Bohn and D. W. Heermann, *Topological interactions between ring polymers: Implications for chromatin loops*, J. Chem. Phys., **132**, 044904, 2010.
102. A. Narros, A. J. Moreno, and C. N. Likos, *Influence of topology on effective potentials: coarse-graining ring polymers*, Soft Matter, **6**, 2435, 2010.
103. J. D. Halverson, W. B. Lee, G. S. Grest, A. Y. Grosberg, and K. Kremer, *Molecular dynamics simulation study of nonconcatenated ring polymers in a melt. I. Statics*, J. Chem. Phys., **134**, 204904, 2011.
104. J. D. Halverson, W. B. Lee, G. S. Grest, A. Y. Grosberg, and K. Kremer, *Molecular dynamics simulation study of nonconcatenated ring polymers in a melt. II. Dynamics*, J. Chem. Phys., **134**, 204905, 2011.
105. T. Sakaue, *Ring polymers in melts and solutions: scaling and crossover*, Phys. Rev. Lett., **106**, 167802, 2011.
106. T. Sakaue, *Statistics and geometrical picture of ring polymer melts and solutions*, Phys. Rev. E, **85**, 021806, 2012.
107. A. Rosa, E. Orlandini, L. Tubiana, and C. Micheletti, *Structure and dynamics of ring polymers: entanglement effects because of solution density and ring topology*, Macromolecules, **44**, 8668, 2011.
108. A. Rosa and R. Everaers, *Ring polymers in the melt state: the physics of crumpling*, Phys. Rev. Lett., **112**, 118302, 2014.
109. D. Michieletto, D. Marenduzzo, E. Orlandini, G. P. Alexander, and M. S. Turner, *Threading dynamics of ring polymers in a gel*, ACS Macro Lett., **3**, 255, 2014.
110. D. Michieletto, D. Marenduzzo, E. Orlandini, G. P. Alexander, and M. S. Turner, *Dynamics of self-threading ring polymers in a gel*, Soft Matter, **10**, 5936, 2014.

111. A. Narros, A. J. Moreno, and C. N. Likos, *Effects of knots on ring polymers in solvents of varying quality*, Macromolecules, **46**, 3654, 2013.
112. M. Bernabei, P. Bacova, A. J. Moreno, A. Narros, and C. N. Likos, *Fluids of semiflexible ring polymers: effective potentials and clustering*, Soft Matter, **9**, 1287, 2013.
113. M. Z. Slimani, P. Bacova, M. Bernabei, A. Narros, C. N. Likos, and A. J. Moreno, *Cluster glasses of semiflexible ring polymers*, ACS Macro Lett., **3**, 611, 2014.
114. A. Rosa, *Topological constraints and chromosome organization in eukariotes: a physical point of view*, Biochem. Soc. Trans., **41**, 612, 2013.
115. A. Rosa and R. Everaers, *Structure and dynamics of interphase chromosomes*, PLoS Comput. Biol., **4**, e1000153, 2008.
116. D. Buck and E. Flapan, *Applications of Knot Theory*, vol. 66 of *Proceedings of Symposia in Applied Mathematics*, American Mathematical Society, 2008.
117. L. Postow, C. D. Hardy, J. Arsuaga, and N. R. Cozzarelli, *Topological domain structure of the Escherichia coli chromosome*, Gene Dev., **18**, 1766, 2004.
118. R. Matthews, A. A. Louis, and C. N. Likos, *Effect of Bending Rigidity on the Knotting of a Polymer under Tension*, ACS Macro Lett., **1**, 1352 – 1356, 2012.
119. P. Poier, C. N. Likos, and R. Matthews, *Influence of rigidity and knot complexity on the knotting of confined polymers*, Macromolecules, **47**, 3394, 2014.
120. A. Takano, Y. Kushida, Y. Ohta, K. Masuoka, and Y. Matsushita, *The second virial coefficients of highly-purified ring polystyrenes in cyclohexane*, Polymer, **50**, 1300, 2009.
121. N. Hirayama, K. Tsurusaki, and T. Deguchi, *Linking probabilities of off-lattice self-avoiding polygons and the effects of excluded volume*, J. Phys. A: Math. Theor., **42**, 105001, 2009.
122. M. Rubinstein and R. H. Colby, *Polymer Physics*, Oxford University Press, Oxford, 2003.
123. J. Roovers, *The melt properties of ring polystyrenes*, Macromolecules, **18**, 1359, 1985.
124. J. Roovers, *Viscoelastic properties of polybutadiene rings*, Macromolecules, **21**, 1517, 1988.
125. G. B. McKenna, G. Hadziioannou, P. Lutz, G. Hild, C. Strazielle, C. Straupe, P. Rempp, and A. J. Kovacs, *Dilute solution characterization of cyclic polystyrene molecules and their zero-shear viscosity in the melt*, Macromolecules, **20**, 498, 1987.
126. D. F. Hodgson and E. J. Amis, *Dilute solution behavior of cyclic and linear polyelectrolytes*, J. Chem. Phys., **95**, 7653, 1991.
127. K. Hur, C. Jeong, R. G. Winkler, N. Lasevic, R. H. Gee, and D. Y. Yoon, *Chain dynamics of ring and linear polyethylene melts from molecular dynamics simulations*, Macromolecules, **44**, 2311, 2011.
128. M. Kapnistos, M. Lang, D. Vlassopoulos, W. Pyckhout-Hinzen, D. Richter, D. Cho, T. Chang, and M. Rubinstein, *Unexpected power-law stress relaxation of entangled ring polymers*, Nat. Mater., **7**, 997, 2008.
129. A. Narros, A. J. Moreno, and C. N. Likos, *Architecture-induced size asymmetry and effective interactions of ring polymers: simulation and theory*, Macromolecules, **46**, 9437, 2013.
130. A. Narros, A. J. Moreno, and C. N. Likos, *Effective interactions of knotted ring polymers*, Biochem. Soc. T., **41**, 630, 2013.
131. E. Kaxiras, *Atomic and Electronic Structure of Solids*, Cambridge University Press, Cambridge, 2003.
132. W. J. Briels, *Transient forces in flowing soft matter*, Soft Matter, **5**, 4401 – 4411, 2009.

Modeling of Solvation Effects for Brownian Dynamics Simulation of Biomolecular Recognition

Neil J. Bruce[1], **Daria B. Kokh**[1], **Musa Ozboyaci**[1,2], **and Rebecca C. Wade**[1,3,4]

[1] Molecular and Cellular Modeling Group, Heidelberg Institute for Theoretical Studies (HITS)
Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany

E-mail: {Neil.Bruce, Daria.Kokh, Musa.Oezboyaci, Rebecca.Wade}@h-its.org

[2] Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS MathComp), Heidelberg University, Im Neuenheimer Feld 368
69120 Heidelberg, Germany

[3] Zentrum für Molekulare Biologie der Universität Heidelberg (ZMBH), DKFZ-ZMBH Alliance
Im Neuenheimer Feld 282, 69120 Heidelberg, Germany

[4] Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University
Im Neuenheimer Feld 368, 69120 Heidelberg, Germany

In Brownian dynamics (BD) simulations, solute particles diffuse in a solvent that is modeled as a continuum that influences interparticle forces, and exerts stochastic and frictional effects on particle motion. BD simulations provide a powerful tool to study biomolecular recognition processes accounting for hydrodynamic interactions. They can, for example, be used to compute bimolecular association rate constants, to dock molecules to predict the structure of their complex, and to investigate the effects of macromolecular crowding on molecular association processes. Biomolecular recognition occurs in an aqueous ionic environment. The first few hydration layers surrounding a biomacromolecule or above an inorganic surface or a lipid membrane show deviations in behaviour from bulk solution. These deviations arise from the particulate nature of water and ions, and the ability of water molecules to make hydrogen bonds and to polarize. The main challenge in constructing continuum solvent models is to not only model the bulk solvent properties but also to provide a reasonable approximation of the effects of the surface and interfacial solvent on biomolecular interactions. Here, we describe the solvation models used in BD simulations to investigate biomolecular recognition, and discuss their advantages and disadvantages and possible future improvements.

1 Introduction

An accurate understanding of how biomolecules interact in solution, both with each other and with small molecules, is of vital importance to the understanding of biological cellular processes. While full atomic detail simulation of all solutes and the aqueous environment of the cell offers a way to investigate these processes, the size and diversity of the system, along with time scales involved presents a problem that is currently computationally intractable. Continuum solvent (CS) simulations of biomolecules offer one approach to reducing the computational complexity.

In this paper, we discuss the use of CS representations in the modeling of molecular interactions and the simulation of biomolecular diffusion. We first introduce the BD method (Section 1.1) and the Simulation of Diffusional Association (SDA) software developed in our group (Section 1.2). In Section 2, we describe various CS models and discuss those used in BD simulations. In Section 3, we consider some of the challenges to using CS

models due to the loss of resolution in a continuum representation, and describe some of the methods developed to meet them. In Section 4, we describe efforts to incorporate hydrodynamic interactions in BD simulations, which are important to correctly model the kinetic behaviour of solutes in solution.

1.1 Brownian Dynamics Simulation

BD simulation is a method to simulate the diffusional motion of solutes with a CS representation. The effects on the interactions between solute atoms of solvent atoms are described by an implicit solvation model and stochastic forces mimic the effects of solvent vibrations due to thermal fluctuation. In BD simulations, a commonly used approach to propagate the trajectory of N particles is to use the algorithm developed by Ermak and McCammon[1] in which the displacement of particle i in a time step $\mathbf{\Delta t}$ is described by

$$\Delta \mathbf{r}_i = \Delta t \sum_{j=1}^{N} \left(\frac{\partial \hat{\mathbf{D}}_{ij}}{\partial \mathbf{r}_j} + \frac{\hat{\mathbf{D}}_{ij}}{k_B T} \cdot \mathbf{F}_i \right) + \mathbf{R}_i \tag{1}$$

where $\hat{\mathbf{D}}_{ij}$ is the 3×3 hydrodynamic diffusion tensor between particles i and j (see Section 4), $\mathbf{F}_i$ is the systematic force acting on particle i and $\mathbf{R}_i$ is a vector describing a random displacement of i, sampled at each simulation step from a Gaussian distribution of mean zero that satisfies the variance-covariance relation

$$\langle \mathbf{R}_i \mathbf{R}_j \rangle = 2 \hat{\mathbf{D}}_{ij} \Delta t \tag{2}$$

for all values of i and j.

If hydrodynamic interactions can be assumed to be negligible, the tensors in Equation 1 where $i \neq j$, and the off-diagonal elements of the $i = j$ tensors, can be set to zero. If the diffusive motion is assumed to be isotropic, the diffusion tensor of each particle i can be replaced by a scalar, and the translational displacement of particle i can be given by

$$\Delta \mathbf{r}_i = \frac{\Delta t}{k_B T} D_i \cdot \mathbf{F}_i + \mathbf{R}_i \tag{3}$$

where $\Delta \mathbf{r}_i$ is the three-dimensional displacement of particle i in the current BD step, D_i is the isotropic infinite dilution translational diffusion coefficient of particle i and $\mathbf{F}_i$ is the force vector acting on i prior to the Brownian dynamics move. The vector $\mathbf{R}_i$ is randomly assigned at each time step, from a Gaussian distribution centred on zero and variance $\langle \mathbf{R}_i^2 \rangle = 6 D_i \Delta t$. Analogous equations can be used to simulate the rotational diffusional motion of the particles.

1.2 Simulation of Diffusional Association (SDA)

The SDA BD software package was originally developed by Gabdoulline and Wade for the calculation of bimolecular association rate constants for protein-protein association.[2,3] The diffusional motion of biomolecules is modeled using rigid molecular conformations and the simplified Ermak-McCammon algorithm given in Equation 3. In addition to rate constant calculations, SDA has also been used to calculate protein electron transfer rates,[4] to predict the structures of protein-protein diffusional encounter complexes[5] and to investigate the interactions of proteins with a gold surface,[6] see Figure 1.

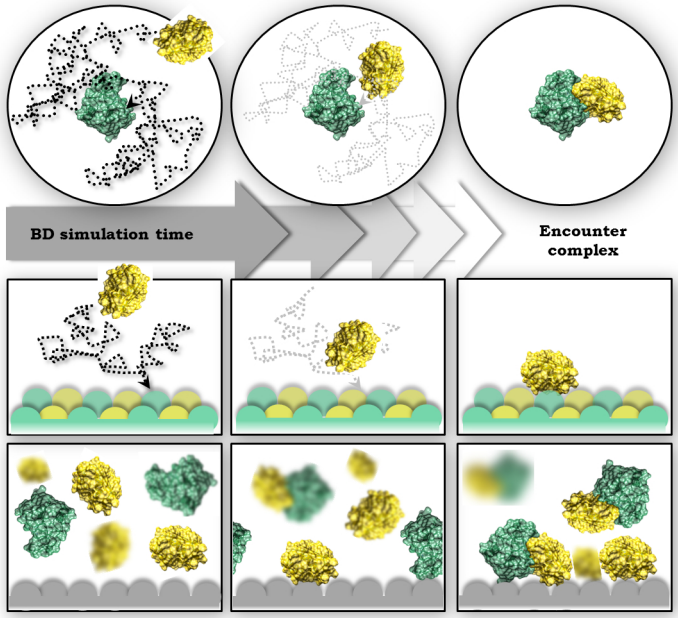


Figure 1: Illustration of systems that can be simulated using SDA. Diffusional association of two macromolecules, such as two proteins, to compute the rate constant for association or the diffusional encounter complex (top); diffusional association of a protein and a surface such as a metal surface (middle); and diffusion and association of many proteins which may take place in the presence of a surface (bottom).

Mereghetti et al[7] extended SDA to allow many macromolecular solute molecules to be simulated in a periodic simulation box. The resulting software, Simulation of the Diffusional Association of Many Molecules (SDAMM), has been used to investigate protein diffusional properties,[7] protein oligomerisation and interactions with a graphite surface,[8] and the effects of macromolecular crowding.[9,10]

More recently, the SDA package for simulating two solutes and the many molecule SDAMM package have been combined into a single distribution, SDA 7, that is available as both a downloadable software package, that includes the main simulation engine and preparation and analysis tools (`http://mcm.h-its.org/sda7/doc/doc_sda7/index.html`), and an interactive webserver (`http://mcm.h-its.org/webSDA`).

2 Continuum Solvent Models

CS models provide computationally efficient and physically intuitive treatments of the effects of solvent that are used in many of types of computation. They have, for example, been applied to compute solute solvation energies, the pK_a values of titratable sites in proteins, and the binding free energies of macromolecular complexes.[11] In BD simulations, the forces between the molecules are computed using CS models. In the next subsection,

we discuss how the different contributions to the solute solvation and binding free energies can be computed. We then describe the CS model used for BD simulations in the SDA software.

2.1 Contributions to Solvation and Binding Free Energies

CS models describe the solvent effects by the sum of polar or electrostatic terms and non-polar terms. The solvation free energy of transferring a solute from vacuum into a solvent at a fixed set of atomic coordinates can be written as

$$\Delta G_{\text{solv}} = \Delta G_{\text{el}} + \Delta G_{\text{np}} \tag{4}$$

where the electrostatic (ΔG_{el}) and non-polar (ΔG_{np}) contributions are approximated by separate and, typically, independent approaches. Together, these terms can provide models that describe phenomena that cannot be described by each alone, e.g. the hydrophobic effect, and they both contribute to the binding free energies between solutes in a solvent. In the following two subsections, a brief overview and discussion of the methods to treat the polar and non-polar terms is given.

2.1.1 Electrostatic Contributions

A range of electrostatic models of varying complexity is available for the electrostatic contribution (ΔG_{el}) to the solvation free energy.[11,12] The simplest models for accounting for solvent in treating interactions between solute charges employ Coulomb's law with a scaled relative dielectric constant. The most common treatment among these models is the empirical scaling of the dielectric constant by making it linearly dependent on the distance between charges, but other scalings, such as sigmoidal,[13] are used as well. These distance-dependent models are used in BD simulations with simple models of solutes, e.g. with each solute represented by a sphere. For simulations in which the spheres represent a protein, the net charge of the molecule can be placed at the centre of the sphere or distributed close to the surface.

In biomolecular simulations, a position-dependent dielectric constant is often employed, where the solutes are considered as low-dielectric cavities ($\epsilon = 1 - 4$) embedded in water represented as a high-dielectric solvent ($\epsilon = 78 - 80$). The most general and accurate treatment of a system of solutes in ionic solution for computing the electrostatic solvation free energy is to solve the Poisson-Boltzmann (PB) equation

$$-\nabla \cdot \epsilon(\mathbf{r}) \nabla \Phi(\mathbf{r}) + \kappa^2(\mathbf{r}) \sinh \Phi(\mathbf{r}) = \rho(\mathbf{r}) \tag{5}$$

where Φ describes the electrostatic potential in an inhomogeneous dielectric medium with a position-dependent dielectric permittivity function, $\epsilon(\mathbf{r})$, and $\rho(\mathbf{r})$ is the density function for the solute charges. The ions in the solvent are treated implicitly and assumed to be distributed around the solutes according to a Boltzmann distribution. The distribution of mobile ions in the solvent are accounted for, typically, with Debye-Hückel theory and included in the equation with κ. The PB equation is a second-order partial differential equation and can be solved analytically only for solutes of simple geometry. For complex systems such as biomolecules, a numerical solution method is required and a variety of finite difference, finite element and multigrid approaches are used. The PB equation can

be linearized to facilitate solution of the equations by assuming that $\sinh \Phi(\mathbf{r}) \approx \Phi(\mathbf{r})$. However, this approximation is not valid for highly charged solutes.

Solving the PB equation requires assigning partial charges in the solute and radii to the solute atoms or particles. The radii are used to compute the dielectric boundary between the solute and the solvent. For atomic detail solute models, the dielectric boundary can be defined using different definitions of the solute surface (see Section 2.1.2). Some treatments smooth the boundary for numerical reasons or to provide a better description of the variation in ϵ in the system. The effect of a sudden change in the ϵ values at the boundary, in particular, can be significant for computed solvation free energies. For proteins, calculations show that the core generally has a lower ϵ than the flexible side-chains on the surface.[14] Therefore, care must be taken in choosing an appropriate solute dielectric constant as well as in treating the dielectric properties of the solute-solvent boundary properly. For a general discussion on dielectric constants and dielectric boundaries, see Ref. 15, and see Section 3.1 for a specific case.

Solution of the PB equation can be computationally demanding and therefore, in BD simulations, approximations to the PB equation are used. The simplest approximation is the test charge model. In this model, the electrostatic potential is computed by solving the PB equation on a grid centred on one molecule (the target molecule). The other solute molecule(s) in the system interact with the target molecule through the forces on their (test) charges as they move over the target molecule's electrostatic potential grid. The low dielectric interior of the target molecule is accounted for by using an appropriate dielectric constant and boundary with the PB equation, whereas that of the other molecule(s) with the test charges is not accounted for. The neglect of the low dielectric interior can be particularly problematic for large macromolecules. This problem can be overcome by using effective charges rather than test charges. In the effective charge model, ECM,[16] effective charges are computed for each solute prior to performing BD simulations. The values of the effective charges are fit to reproduce the electrostatic potential computed by solution of the PB equation (using a heterogeneous dielectric) in a shell region around the solute surface using a uniform dielectric constant. During the BD simulations, forces between molecules are computed using the effective charges in the same way as the test charges in the test charge model. The effective charge model has been widely used. McGuffee and Elcock[17] employed the ECM method in a BD simulation protocol to simulate 1000-molecule systems with all-atom models. Similar to ECM, the DiSCO (discrete surface charge optimization) procedure[18] was developed in order to reproduce electrostatic potentials obtained by the non-linear PB equation using effective Debye-Hückel charges distributed on a virtual surface around the biomolecule. Beard & Schlick applied their method to BD simulations of nucleosome models with electrostatic interactions to study their condensation.[19]

Noteworthy is that the ECM model can be used for computing intermolecular forces and energies but is not applicable to intramolecular forces in simulations with flexible molecules. Thus, in programs like SDA (see Section 2.2) and Browndye[20], the simulations employ rigid-body models of the solutes (though SDA can describe the flexibility of a molecule by switching periodically between a set of rigid conformations). Other BD packages, such as BD_BOX,[21] SIMUFLEX[22] and Brownmove,[23] developed for flexible simulation of coarse-grained models of biomolecules, make use of a screened Coulomb or Debye-Hückel approximation for electrostatic interactions.

A number of CS models for electrostatic interactions of intermediate complexity be-

tween a simple Coulomb interaction and solution of the PB equation have been developed. The dielectric variation between the solute and solvent regions can be accounted for by computing the burial of the charges in the solute and using this for a rescaling of the dielectric constant in Coulomb's law or using a method of image formalism assuming a planar boundary between the solute and the solvent.[24] A commonly used empirical alternative to the PB-based methods is the Generalized Born (GB) model which includes the self-energy contribution from individual atoms and an atom-pair correction.[25,26] The GB approximation extends the spherical Born model to cavities of complex shapes. Although less accurate, the GB method is computationally cheaper than solving the PB equation and, therefore, commonly used in implicit solvent MD simulations. As with the PB equation, solution of the GB model depends on empirical parameters, such as the effective radii of ions. Such intermediate CS models could be applicable in BD simulations of biomacromolecules.

2.1.2 Non-polar contributions

The non-polar solvation free energy, ΔG_{np}, can be considered to consist of three major energetic terms

$$\Delta G_{\text{np}} = \Delta G_{\text{cav}} + \Delta G_{\text{arr}} + \Delta G_{\text{att}} \tag{6}$$

The first and the second terms in the equation above describe the energetic costs, respectively, of creating a cavity in solvent to accommodate a solute molecule and of rearranging the water molecules at the solute-solvent interface. Together, they represent the repulsive contributions from solute-solvent interactions to the overall solvation free energy. The third term, on the other hand, describes the attractive van der Waals dispersion interactions between the solute and the solvent. In the simplest models of ΔG_{np}, it can be calculated using a single surface area or volume-dependent term. Alternatively, the cavity and interface contributions to ΔG_{np} can be estimated by simple volume and solvent accessible surface area (SASA) dependent functions, respectively, and the van der Waals interactions estimated by surface integral-based methods.[27–29] A number of different schemes exist for separating ΔG_{np} into attractive dispersive terms and repulsive terms.[30] Full models to calculate non-polar solvation energies by computing each of the terms individually[31,32] have been shown to be more accurate in reproducing the explicit solvent solvation than those that approximate the total free energy by SASA (only or with volume) dependent functions. In order for mean-field and other approximations in these models to reproduce the overall non-polar solvation forces and energies, parameters such as surface tension and probe radius to define the SASA of a solute need to be parameterized. Therefore, applications of the models to different types of molecules require extensive calibration and testing. One of the full models developed[32] demonstrated that, while the model reproduced the mean non-polar solvation forces accurately for a protein in solvent, it failed to estimate solvation energies of small alkane solutes correctly, necessitating readjustment of the solvent probe radius to 0.65 Å, which is much smaller than the traditional 1.4 Å water probe radius, in order to match the experimental solvation energies of these molecules. In cases when they are accurate, the models developed for dilute solute-solvent systems fall short in calculation of non-polar desolvation free energy of macromolecules forming complexes, since perturbations of the structure and changes in the density of the interfacial solvent due to a

second solute cannot be accurately approximated by mean field methods. In many cases, a simple surface-area dependent term to describe non-polar contributions to binding is employed. The parameterisation of this term depends on the definition of the surface used: the SASA of an atomic detail solute (defined by the centre of an interacting probe solvent sphere) is greater than the molecular surface (MS) of a solute (defined by the contact and reentrant surfaces of an interacting probe solvent sphere). Such a SASA-based non-polar desolvation free energy term is employed in SDA (see Section 2.2).

The effective boundary definition between the solute and the solvent depends on non-polar as well as electrostatic interaction energy terms. However, the solute-solvent boundaries are usually defined only for the term they are intended to be used for. To address this issue, the Variational Implicit-Solvent Model (VISM) was developed for coupling the electrostatic and non-polar energy terms.[33–35] In this approach, the Gibbs free energy of the solvent is minimized with respect to a function for solvent volume exclusion, giving the solvent-accessible surface of the solute as output. Although VISM offers some advantages over other commonly used methods, such as boundary selection based on radii parameters, it requires further calibration. Furthermore, it is very costly to solve the PB equation for each of the optimization steps and therefore applications of VISM are currently limited.

2.2 Treatment of the Effect of Solvent on Interaction Energies and Forces in SDA

Using the PB equation to calculate the electrostatic interactions between all solutes at each step of a BD simulation is costly. To reduce the computational cost of calculating these interactions, SDA exploits the fact that the electrostatic field produced by a rigid solute can be approximated as constant throughout a trajectory. This allows it to be computed prior to simulation, and for the interaction at each simulation step to be modeled as the interaction of a set of effective charges (see Section 2.1.1) on one solute with a constant field surrounding its interacting partner.

The electrostatic model used in SDA was described by Gabdoulline and Wade,[16] and was first applied to study the diffusional association of the extracellular ribonuclease barnase with its intracellular inhibitor barstar.[2,3] Three-dimensional models of the solutes are generated with appropriate protonation states of the titratable groups for the experimental conditions of interest. Atomic partial charges and radii are assigned from a suitable force field and the PB equation is solved for each solute using an external PB solver to generate the molecular electrostatic potential grids, $\mathbf{\Phi}_{\text{el}}(\mathbf{r}_i)$.

During simulations, the electrostatic interaction energy between a pair of solutes is modeled as the interaction of effective charges (see Section 2.1.1) on one solute with the potential field of another, according to

$$\Delta G_{\text{el}}^{1-2} = \frac{1}{2}\sum_{i_1} q_{i_1}\, \mathbf{\Phi}_{\text{el}_2}(\mathbf{r}_{i_1}) + \frac{1}{2}\sum_{i_2} q_{i_2}\, \mathbf{\Phi}_{\text{el}_1}(\mathbf{r}_{i_2}) \tag{7}$$

where $\mathbf{\Phi}_{\text{el}_n}(\mathbf{r}_{i_m})$ is the electrostatic potential of solute n at the position of charge site i on solute m. The factor of $\frac{1}{2}$ is required to avoid double-counting of the two-body potential.

The long range inverse distance dependence of electrostatic interactions means that the precomputed potential grids need to be sufficiently large that the electrostatic potentials at the grid boundaries are so small that truncation errors are negligible. This can require the use of large grids that create a significant memory burden on the calculations. This is

particularly a problem during simulations at low ionic strength conditions, where the low ionic screening means that the potential decays slowly with distance. Mereghetti et al[36] introduced a long range Debye-Hückel correction into SDA to address this problem, while removing the need for large grid sizes. At small separations, electrostatic interactions are modeled using the grid-based approach described above. At larger separations, where the electrostatic field surrounding a solute is approximately spherically symmetric, electrostatics are modeled as the interaction of a pair of Debye-Hückel spheres with charges z equal to their net effective charges and radii a equal to their Stokes hydrodynamic radii. The resulting correction takes the form

$$\Delta G_{\text{DH}}^{1-2}(r) = \frac{z_1 z_2 e_l^2 e^{-\kappa(r-(a_1+a_2))}}{4\pi\epsilon_0\epsilon_s r(1+\kappa(a_1+a_2))} \tag{8}$$

where e_l is the elementary charge, ϵ_0 and ϵ_s are the permittivity of vacuum and the solvent relative dielectric constant, respectively, and κ is the inverse Debye length, which is dependent on the solvent ionic strength.

When two diffusing solutes approach each other closely, the cost of desolvating the charges on the solutes should be accounted for. This is done in the approximation to the PB forces used in SDA by computing an electrostatic desolvation term describing the interaction between point charges on one solute and low dielectric spherical cavities in the interacting solute.[16,37] This term is again computed using precomputed grids, the electrostatic desolvation contribution to the free energy of interaction between a pair of solutes given by

$$\Delta G_{\text{edesolv}}^{1-2} = \sum_{i_1} q_{i_1}^2 \, \mathbf{\Phi}_{\text{edesolv}_2}(\mathbf{r}_{i_1}) + \sum_{i_2} q_{i_2}^2 \, \mathbf{\Phi}_{\text{edesolv}_1}(\mathbf{r}_{i_2}) \tag{9}$$

where q_{i_m} is an effective charge on solute m and $\mathbf{\Phi}_{\text{edesolv}_n}(\mathbf{r}_{i_m})$ is the electrostatic desolvation potential of solute n at the position of this effective charge, defined at a point $\mathbf{r}$ by

$$\mathbf{\Phi}_{\text{edesolv}}(\mathbf{r}) = \alpha \frac{\epsilon_s - \epsilon_p}{\epsilon_s(2\epsilon_s + \epsilon_p)} \sum_j a_j^3 \frac{(1+\kappa r_j)^2}{r_j^4} e^{-2\kappa r_j} \tag{10}$$

where ϵ_s and ϵ_p are the solvent and solute dielectric constants, respectively, κ is the inverse of the Debye length and α is an empirical parameter used to scale the interaction strength. The sum in Equation 10 runs over all solute atoms j of radius a_j and r_j is the distance from atom j to point $\mathbf{r}$.

The reduction in the total solute/solvent interface when two solutes form a binding interface results in a favourable non-polar contribution to their interaction free energy. In SDA, this non-polar interaction is modeled by using a surface-area dependent term given by[4]

$$\Delta G_{\text{np}}^{1-2} = \sum_{i_1} \text{SASA}_{i_1} \, \mathbf{\Phi}_{\text{np}_2}(\mathbf{r}_{i_1}) + \sum_{i_2} \text{SASA}_{i_2} \, \mathbf{\Phi}_{\text{np}_1}(\mathbf{r}_{i_2}) \tag{11}$$

where SASA_{i_m} is the solvent accessible surface of atom i in solute m and $\mathbf{\Phi}_{\text{np}_n}(\mathbf{r}_{i_m})$ is the non-polar desolvation potential of solute n at the position of atom i in solute m. The

non-polar desolvation potential at a point $\mathbf{r}$ is given by the conditional expression

$$\mathbf{\Phi}_{\text{np}}(\mathbf{r}) = \beta c \begin{cases} 1 & r_{\text{min}} < a \\ \frac{b - r_{\text{min}}}{b - a} & a < r_{\text{min}} < b \\ 0 & r_{\text{min}} > b \end{cases} \tag{12}$$

where r_{min} is the minimum distance from point $\mathbf{r}$ to an atom on the surface of the solute, while a describes the maximum distance from the solute surface at which a probe atom is considered to be completely desolvated by the solute and b the minimum distance a probe atom can be from the solute surface while remaining completely solvated. Parameter β is the constant of proportionality between the total change in surface area and the change in interaction free energy, and parameter c is present to normalize the sum in Equation 11.

The repulsive interactions that prevent solutes from overlapping are modeled in two ways in SDA. In the first method,[2,3] an excluded volume check is performed following every BD move to check that no atoms in one solute have overlapped with the solvent accessible surface of another. If an overlap is found, the BD move is repeated with a different random displacement until a configuration with no overlap is created. This overlap check can be expensive when simulating the dynamics of many solutes, so it can be replaced with a soft-core repulsive term,[7] that is calculated using precomputed grids with

$$\Delta G_{\text{rep}}^{1-2} = \sum_{i_1} E_{\text{rep}_2}(\mathbf{r}_{i_1}) + \sum_{i_2} E_{\text{rep}_1}(\mathbf{r}_{i_2}) \tag{13}$$

where $E_{\text{rep}_n}(\mathbf{r}_{i_m})$ is the soft-core repulsive potential of solute n at the position of surface atom i_m of solute m. The soft-core repulsive potential surrounding a solute with surface atoms i, radius a_i, is given by

$$E_{\text{rep}}(\mathbf{r}) = \gamma \sum_i \frac{1}{(\frac{a_i}{\sigma})^{\text{nexp}} + |\mathbf{r} - \mathbf{r}_i^{\text{nexp}}|} \tag{14}$$

where γ, σ and nexp are suitable parameters.[7]

3 Challenges Due to the Particulate Nature of Water and Ions

CS models provide a good balance between accuracy and computational efficiency and are, therefore, commonly used in computer simulations of biomolecules. However, the computational efficiency comes at the cost of falling short in reproducing the effects of explicit solvent as the properties of the water at solute-solvent interfaces deviate significantly from those of bulk solvent.[38] Current implicit solvent models take into account the screening effect of the bulk solvent and hydrophobic solvation effects, leaving the permittivity and granularity of interfacial solvent atoms that cannot be described by the macroscopic bulk properties out of the equation. These microscopic features contribute to the structure and the dynamics of the interfacial solute atoms and mediate biomolecular association. In this section, we will discuss the significance of certain local solvent properties and their impact in CS methods.

3.1 Relative Permittivity of the Water

Different implicit solvation models treat the effect of the solvent permittivity differently and this effect is included with distinct dielectric screening functions. Among these solvation models, PB and GB commonly employ a position-dependent dielectric function that separates the solute and solvent with a boundary and assigns distinct dielectric constants within each region to account for the permittivity of the overall system. The internal (solute) and the external (solvent) dielectric constants, as well as the definition of the boundary that separates these two regions, are empirical parameters per se that are only defined in the context of the full model.[39,40] To date, internal dielectric constants for proteins varying from 1, 2 and 4[41] up to 20[42] and even higher (close to the solvent dielectric) for pK_a calculations of solvent exposed residues[43] have been used together with several different dielectric boundary definitions such as molecular as well as van der Waals and solvent accessible surfaces of the solute.[44,45] Further, computed solvation energies are highly sensitive to the dielectric boundary definition and the dielectric constant used in different approaches[46] and therefore every new dielectric function requires a new parameterisation.[47]

The force fields commonly used in molecular simulations have been developed and calibrated for the accurate representation of solute-solute and solute-solvent interactions in the presence of an explicit solvent medium. Using parameters such as partial atomic charges and Lennard-Jones radii from these force fields with implicit solvent models can introduce incompatibilities and therefore the parameters may require adjustment.[48] To account for the implicit solvation correctly, optimized atomic radii parameters[49] were proposed and the PARSE force field[50] was developed for the GB and PB models respectively. Even though these optimized parameters have proved to be successful in the reproduction of the explicit solvation effects on average, they still fall short in describing these effects accurately especially when the structure and dynamics of the water at a microscopic level play a prominent role in mediating certain molecular events, e.g. adsorption.

As a solvent medium, water is highly polarisable and has an experimentally observed bulk relative dielectric constant of about 78 at room temperature. When two solutes solvated in water are far apart from each other, the electrostatic interactions between them are screened and this can be explained simply by the dielectric constant of the bulk water. However, when these solutes are close enough that the structure of the interfacial water deviates significantly from that of the bulk water, the electrostatic screening between them cannot be explained simply by a constant bulk relative dielectric constant. The same applies to any confined water in a system. In a recent study, it was found that not only does the dielectric constant of the water within a few nanometres of the interfaces between hydrophilic proteins deviate significantly from that of bulk but that it is also anisotropic in nature.[51] The permittivity of the water, therefore, can be represented not by a constant but by a tensor. In the same study, it was shown from molecular dynamics simulations of protein-protein association in explicit solvent that the value of the dielectric tensor in the direction of association is lower than those in the other two orthogonal directions, thus leading to a stronger effect of the electrostatic interactions on association. This behaviour was attributed to an adhesive hydrogen bond network at the interface that was observed in the simulations.

We investigated the role of the water and its permittivity for the adsorption of the beta-lactamase inhibitor protein (BLIP) tagged with three histidine residues on its N terminus (3H-BLIP) on an Au(111) surface in molecular dynamics simulations with explicit water

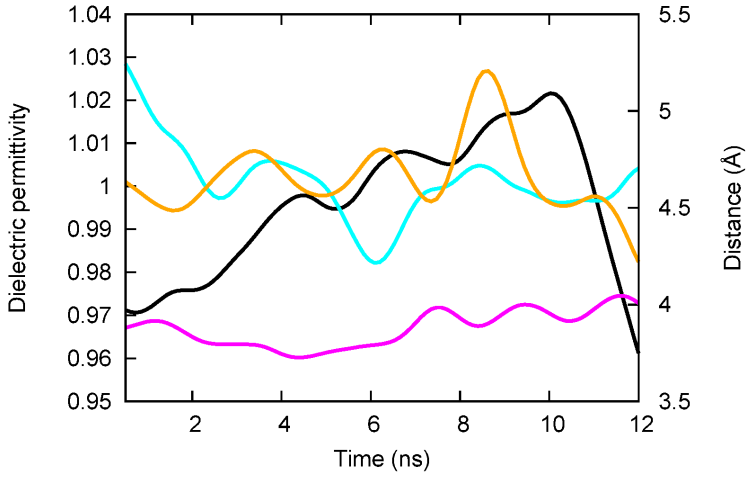


Figure 2: Relative dielectric permittivity tensor of the interfacial water along the x (cyan), y (orange) and z (magenta) axes and the distance between the 3HIS tag of the 3H-BLIP protein and an Au(111) surface in the z-direction (black). The values of the permittivity tensor are normalized relative to the values for bulk water. (MO, DBK, RCW, unpublished data)

(unpublished data). The simulations showed that water forms three structured layers on the gold surface. In order for a solute molecule to adsorb on the gold surface, these layers of water molecules that are structurally different than that of bulk should be displaced. Further, using the data from the simulations, we computed the directional permittivities. To this end, we calculated the Cartesian components of the dielectric tensor of the water in the binding interface using the equation given by Ahmad et al.[51]

$$\frac{(\epsilon-1)(2\epsilon+1)}{(3\epsilon)} = \frac{4\pi}{Vk_BT}\left(\left\langle(\mathbf{M}\cdot\mathbf{e})^2\right\rangle_{\mathbf{E_d}} - \langle\mathbf{M}\cdot\mathbf{e}\rangle^2_{\mathbf{E_d}}\right) \tag{15}$$

where ϵ is the relative permittivity; V and k_BT are the volume of the sample and product of the Boltzmann constant and the temperature, respectively; and $\mathbf{M}$, $\mathbf{e}$ and $\mathbf{E_d}$ are the total dipole moment of the sample, a unit vector and the directing field, respectively. The gold surface was located on the x-y plane and the adsorption of 3H-BLIP occurred by reduction of the z coordinate of the protein. As shown in Figure 2, the permittivity is clearly anisotropic. The components in the x and y directions fluctuate around the dielectric constant of the bulk solvent, whereas the component in the z direction is distinctly lower. The weaker screening in the z direction hence increases the strength of the electrostatic interactions between the metal surface and the protein, mediating the adsorption process.

3.2 Effect of Hydrogen Bonds

The issue of granularity becomes even more significant for the implicit representation of the solvent in the direct vicinity of the solute-solvent interface. Water forms hydrogen bonds with the solute atoms and they are not explicitly represented in CS models. A

correction term at a mean-field level is added to the overall solvation energy to account for the hydrogen bonding interactions in these models.[52,53,39,54] However, this correction does not account for the interactions between solute molecules mediated by water through hydrogen bonds.

Studies show that the interfacial water not only plays a role in effective screening of electrostatic interactions but also in mediating interactions between polar groups on solute surfaces by forming water bridges.[51,55] Ahmad et al (2011) found, by employing a maximum flow analysis, that the interactions between hydrophilic surface patches of proteins are mediated and stabilized by stable hydrogen bonding networks starting at early transient encounter phases. These indirect interactions arising from the granularity of the solvent molecules and therefore need to be modeled in future continuum solvation models.

In conclusion, CS models have yet to reproduce all aspects of solvent environments completely, in particular of water near a membrane bilayer[56] or ion channels,[57] in addition to those discussed earlier.[38] To overcome problems that arise due to the mean-field level treatment of certain microscopic features, hybrid approaches have been proposed, such as explicit representation of the hydration shells around biomolecules[58,59] and ions[60] with continuum solvation applied to the rest of the system. In addition to hybrid approaches, coarse-grained models of water for simulations of biomolecular systems that capture some of aspects of all atom explicit water models, for instance entropy and even the permittivity of bulk water, have been proposed.[61,62] However, they also fall short in reproducing the microscopic properties of water at solute interfaces due to their exclusion of hydrogen-bonding interactions from the models and larger size. Thus, for BD simulations, there is a need to develop CS models to account for deviations from bulk behaviour near solute surfaces. One example of such a CS model is given in the next section.

3.3 Example: Metal Desolvation by a Protein

A CS model, including a short-range desolvation correction to the total interaction energy, as described Section 2.2, usually provides an adequate level of accuracy for rigid body BD simulations of protein association. In the case of protein adsorption to a metal or inorganic surface, however, long-range electrostatic interactions may not be dominating. Indeed, adsorption to a hydrophobic surface (such as graphite or a surface functionalized by a self-assembled monolayer of hydrophobic compounds) is driven by pure short-range interactions. The electrostatic interaction with metal surfaces arises solely from polarization effects and is thus much weaker than for protein-protein binding. Also, protein electrostatic interactions with a hydrophilic surface rapidly decrease with the protein-surface distance if a biomolecule has a small net charge, because the effects from different partial charges in the protein cancel out.

On the other hand, short-range interactions from van der Waals forces, charge transfer, and distortion of the hydration shell, often appear to be stronger than for macromolecular interactions. Since many hydrogen bonds are broken and water builds an extended ordered layer, hydration shell properties, in particular variations of physical (such as oscillation of density, free energy, and electrostatics) and chemical characteristics (surface reconstruction, ionization, hydration, etc) are different from those of a biomolecule.[63] A typical example of the water density in a hydration shell is shown in Figure 3a. Bound water in the first hydration layer on the gold surface is separated by a free energy barrier

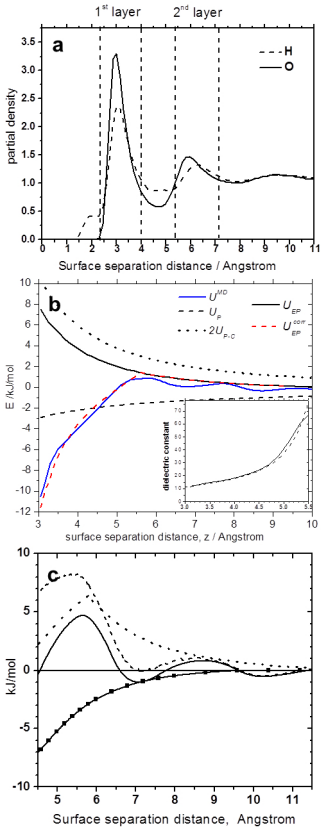


Figure 3: Modeling of short-range corrections to a CS force field, ProMetCS, for simulation of protein adsorption on a Au(111) surface (Reproduced with permission from Ref 6. Copyright (2010) American Chemical Society). (a) Dependence of the partial density of the water oxygen atoms (solid line) and hydrogen atoms (dashed line) on the distance from the gold surface computed from MD simulations of water in the presence of a gold surface. Densities are normalized to the bulk values. (b) Total electrostatic energy for a test charge atom as a function of distance from the gold surface with explicit (U^{MD}) and implicit ProMetCS ($U_{\mathrm{EP}}^{\mathrm{corr}}$) water models. U_{P}, U_{EP}, and $U_{\mathrm{P\text{-}C}}$ are separate CS contributions to the electrostatic energy: image-charge energy of an ion in continuum solvent with a dielectric constant of 80, electrostatic image charge energy with an electrostatic desolvation term as used in the ECM model (see Section 2.2), pure electrostatic desolvation term taking into account the low-dielectric layer above the metal surface and also on the protein surface, respectively. Insert: Plot of effective dielectric constant derived from the image-charge potential computed from explicit water simulations (solid line) and approximated by an analytical function (dashed line). (c) Potential of mean force (PMF) for a test atom as a function of atom-gold surface distance, as obtained from MD simulations (solid line), corresponding Lennard-Jones potential (squares), and their difference (dashed lines) associated with the desolvation energy. Dotted line - PMF energy computed using the ProMetCS model (includes both Lennard-Jones and metal hydrophobic desolvation energies, see text).

of $\sim$4.4 kJ mol^{-1} from the bulk water, which gives rise to an unfavorable positive energy change of the solvent-metal system when a water molecule is removed from the hydration shell to the bulk. The properties of the hydration shell depend strongly on the chemistry

of the surface. Also the adsorption of ions and other small molecular species varies considerably for different solids; for example, for the gold surface it was found that citrate (in comparison with chloride) ions were preferably adsorbed.[64] A review of studies of water properties covering a wide range of the surfaces can be found elsewhere.[65]

The standard implicit solvent models often fail to adequately describe recognition processes and to achieve even qualitatively the characteristics of protein adsorption (such as an identification of the protein adsorption site). Therefore, a more detailed description of the short-range interactions, and the adaptation of the implicit solvent model to each type of surface, is generally required. An example of such a force field is ProMetCS, a protein-metal continuum solvent force field that has been parameterized for Au(111) surfaces.[6] In ProMetCS, the protein-surface interactions are described in atomic detail using the force field parameters of the Lennard-Jones interaction term between organic molecules and a solid state surface as parameterized by Iori et al.[66] Metal polarization above the hydration shell is represented by an image-charge model using an effective charge approximation similar to the protein-protein electrostatic model in a continuum medium with a fixed dielectric constant. However, the electrostatic interactions of the protein effective charges placed at distances of less than 0.6 nm from the metal surface (i.e. within the first two surface water layers) are computed directly as the Coulomb interaction energy between an effective charge and its image using a distance-dependent dielectric constant. Dielectric constant variation with the distance from the surface was parameterized using the potential of mean force of an ion on the Au(111) surface (Figure 3b) obtained from MD simulations taking polarization effects into account using a Druid model.[67] This term ensures strong electrostatic binding of charged residues as soon as they penetrate the hydration shell. It was shown in MD simulations that the short-range electrostatic effect also plays an important role in the adsorption process on charged surfaces.[68]

The surface desolvation free energy arising from the partial replacement of the metal hydration shell by a protein adsorption site is described in ProMetCS by a function that is proportional to the protein-metal contact area. A free energy change for desolvation of a unit surface area of the metal was derived from the PMF of the test atom obtained from MD simulations (Figure 3b, along with the corresponding van der Waals interaction energy). Since the translational entropy change along the PMF is zero for the present case, the difference between the PMF and LJ energies (dashed line in Figure 3c) corresponds to the metal desolvation energy. It shows maxima at the first and second hydration layers at surface separation distances of $\sim$5.5 Å and $\sim$8.5 Å, respectively. In ProMetCS, the surface desolvation energy is assumed to be constant within the first hydration layer and then decrease exponentially within the second layer with a cutoff of 10 Å. The PMF of an atom on the gold surface computed using the ProMetCS force field demonstrates the smoothed profile of the PMF function obtained in explicit solvent MD simulations (Figure 3c).

4 Hydrodynamic Interactions

As a solute moves in solution, it creates a flow current within the solvent that couples its dynamics to those of all other solutes. When performing biomolecular simulations with an explicit solvent representation, these currents, known as hydrodynamic interactions (HI) are explicitly modeled by the motions of the solvent. In a standard CS model, they are absent, and must be modeled explicitly into the simulation dynamics, if their effects are to

be included.

In dilute solutions of biomacromolecules, the effects of intermolecular HI are often, to a first approximation, assumed to be negligible. However, the validity of this approximation needs to be carefully considered. Indeed, accounting for HI in BD simulation has been suggested to be important in studies of bimolecular association,[69] protein folding[70,71] and membrane formation.[72]

The Ermak-McCammon BD algorithm,[1] described in Section 1.1, implicitly includes HI. However, the apparent simplicity of the algorithm, as shown in Equation 1 (Section 1.1), hides a great degree of computational complexity. At all simulation time steps, the pair hydrodynamic diffusion tensors $\hat{\mathbf{D}}_{ij}$ must be recalculated, and then factorized according to Equation 2 to generate the random fluctuations. A number of schemes have been reported to approximate the tensors $\hat{\mathbf{D}}_{ij}$ and perform the subsequent factorisation.

4.1 Brownian Dynamics with Hydrodynamic Interactions

Much of the early work in modeling diffusion was in the study of polymer chains represented by simple bead models, and one of the most commonly used descriptions of the hydrodynamic tensors comes from this field. Rotne and Prager[73] derived an approximation for the hydrodynamic diffusion tensors of equally-sized spheres. Their resulting tensors are built on the earlier work by Oseen as described by Kirkwood,[74] who treated the diffusing particles as frictional points, an approximation that was found to be inadequate at small separations. Rotne and Prager's approximation was also derived independently by Yamakawa,[75] and developed further by De La Torre and Bloomfield[76] to account for differing sphere sizes. The hydrodynamic tensors, when $i \neq j$, are given by

$$\hat{\mathbf{D}}_{ij} = \frac{k_B T}{8\pi\eta r_{ij}} \left(\hat{\mathbf{I}} + \frac{\mathbf{r}_{ij}\mathbf{r}_{ij}}{r_{ij}^2} + \frac{a_i^2 + a_j^2}{r_{ij}^2} \left(\frac{1}{3}\hat{\mathbf{I}} - \frac{\mathbf{r}_{ij}\mathbf{r}_{ij}}{r_{ij}^2} \right) \right) \tag{16}$$

where k_B is the Boltzmann constant, T is temperature, η is the solvent viscosity, $\hat{\mathbf{I}}$ the 3×3 unit tensor, a_i and a_j the radii of particles i and j, respectively, and $\mathbf{r}_{ij}$ the vector of length r_{ij} from particle i to particle j. Equation 16 is only valid in the case of non-overlapping particles (i.e. when $r_{ij} > a_i + a_j$). For overlapping particles, an alternative equation is used

$$\hat{\mathbf{D}}_{ij} = \frac{k_B T}{6\pi\eta r_{ij}} \left(\left(1 + \frac{9}{16}\frac{r_{ij}}{(a_i + a_j)} \right) \hat{\mathbf{I}} + \frac{3}{16}\frac{\mathbf{r}_{ij}\mathbf{r}_{ij}}{r_{ij}(a_i + a_j)} \right) \tag{17}$$

While strictly only applicable when $a_i = a_j$, Equation 17 has also been applied in the case where $a_i \neq a_j$.[21] When $i = j$, the hydrodynamic tensor is given by

$$\hat{\mathbf{D}}_{ij} = \frac{k_B T}{6\pi\eta r_{ij}} \hat{\mathbf{I}} \tag{18}$$

Ermak and McCammon[1] showed that the random fluctuations in Equation 1, that satisfy Equation 2, could be obtained by factorising the $N \times N$ diffusion tensor $\hat{\mathbf{D}}$, whose elements are the 3×3 hydrodynamic pair tensors $\hat{\mathbf{D}}_{ij}$, by Cholesky decomposition

$$\hat{\mathbf{D}} = \hat{\mathbf{S}}\hat{\mathbf{S}} \tag{19}$$

The $3N$ fluctuation vector $\mathbf{R}$ that comprises the $\mathbf{R}_i$ vectors of Equation 1 for each particle i can then be obtained from

$$\mathbf{R} = \hat{\mathbf{S}}\mathbf{X} \tag{20}$$

where $\mathbf{X}$ is a vector of length $3N$, whose elements are Gaussian distributed random numbers, of mean zero and variance $2\Delta t$. Calculating these fluctuations by factorising $\hat{\mathbf{D}}$ formally scales as N^3, and is the most expensive part of a BD simulation that incorporates HI.

To improve the performance of BD simulations, a number of approaches have been developed. One of the first such approaches is the Chebyshev polynomial approximation of Fixman,[77] which scales as $N^{2.5}$. Geyer and Winter described their truncated expansion ansatz that models HI in BD simulations (TEA-HI).[78] In this method, which scales as N^2, HI are approximated though a first-order expansion when calculating the random fluctuations in the Ermak-McCammon equation while using the full tensor in the force term. This approach has been included in the BD package Brownmove.[23] In the BD_BOX simulation package,[21] the use of GPUs to perform HI calculations is supported. This allows fast factorisation of the hydrodynamic diffusion tensor, using either Cholesky decomposition or Fixman's Chebyshev polynomial approximation.

Elcock[79] took a different approach to speeding up HI simulations of flexible molecules by separating the intramolecular and intermolecular HI terms. Intramolecular HI were modeled using the full diffusion tensor, while intermolecular HI were modeled by assuming that the solutes were spheres. Using coarse-grained protein models and upto 1024 proteins per BD simulation, he observed improved computational performance compared to the full-HI approach.

4.2 Mean-field Approximation of Hydrodynamic Interactions

An alternative method of including HI in BD simulations is through the mean-field approach first described by Heyes.[80] This method is included in SDA, as described by Mereghetti et al.[10] The correlated nature of HI are neglected, and instead the simplified version of the Ermak-McCammon algorithm described in Equation 3 is employed, removing the need for the expensive factorisation of the hydrodynamic diffusion tensor. HI are reincorportated into the dynamics of each solute by scaling the infinite dilution translational and rotational diffusion coefficients, D_i^{t} and D_i^{r}, used in Equation 3 according to the local volume fraction, ϕ_i, surrounding the solute i that is occupied by solutes in a spherical volume of radius R_{cut} centred on the solute i (Figure 4a).

The local volume fraction is calculated (Figure 4b) as

$$\phi_i = \frac{V_i + \sum_j V_j + \sum_k V_k}{\frac{4}{3}\pi R_{\mathrm{cut}}^3} \tag{21}$$

where V_i is the volume of solute i, V_j is the volume of solute j that is completely included in the local volume, and V_k is the partial volume of solute k that is included in the local volume. For mean-field HI calculations, the solutes are considered as spheres with radii equal to their Stokes hydrodynamic radii, and the partial volumes V_k are calculated from the union of the sphere centred on solute k with the Stokes radius of solute k and the sphere of radius R_{cut} centred on solute i.

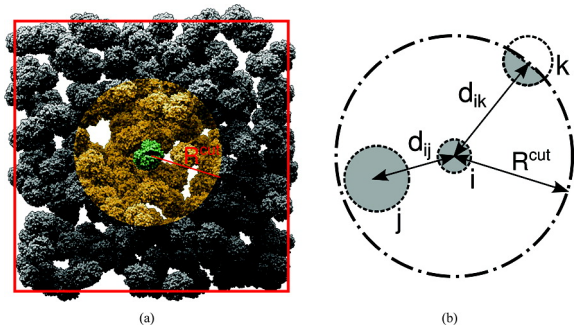


Figure 4: Schematic representation of a simulation box of protein molecules and the model used for the calculation of local volume fraction. (a) As an example of a simulation box, we show here a simulation snapshot of a Haemoglobin A solution at 0.11 volume fraction. Taking as representative the green protein in the centre, the spherical volume of radius R_{cut} used to compute the local volume fraction is indicated in yellow. (b) For the calculation of the local volume fraction, proteins, which can be of different sizes, are described with their hydrodynamic radius (Reproduced with permission from Ref. 10. Copyright (2012) American Chemical Society).

Mereghetti et al[10] showed that the inclusion of mean-field HI improved the agreement between the computed translational self-diffusion coefficients of myoglobin and haemoglobin A and experimental values in high concentration solutions, compared to simulations that neglected HI.

5 Concluding Remarks

The computational efficiency of BD simulations arises in large part due to the use of CS models, which result in systems with far fewer particles than those modeled with explicit solvent models. CS models are well suited to studying biomolecular recognition processes guided by long-range forces such as electrostatic interactions. Thus BD simulations can give reliable results for computations of the effects of changes in pH or ionic strength or of charge mutations on diffusion-influenced bimolecular association kinetics. For recognition processes in which short-range forces, such as hydrogen bonding and hydrophobic interactions, play an important role, CS models can be designed to capture the most important effects of the solvent's particulate nature. For example, in the ProMetCS model for protein-metal interactions, the cost of a solute displacing water from the first two layers of water above the surface is explicitly accounted for in the CS model. Such additions to CS models can be made with little added computational cost. However, there is scope for improving CS models for simulating biomolecular recognition. More advanced multiresolution and multiscale solvation models can be envisaged, as well as advances in models of hydrodynamic interactions.

Acknowledgments

The authors thank all the current and former members of the Wade group that have contributed over the last two decades to the development of the SDA software and gratefully acknowledge the support of Heidelberg Institute for Theoretical Studies (HITS), the

Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS MathComp), the Klaus Tschira Foundation (KTS), the German Federal Ministry of Education and Research (BMBF) Virtual Liver Network (grant no. 0315749), the EU/EFPIA Innovative Medicines Initiative (IMI) Joint Undertaking, K4DD (grant no. 115366), and the EU FEP Flagship Programme Human Brain Project (grant no. 604102). This chapter reflects only the authors' views and neither the BMBF, the IMI nor the European Commission is liable for any use that may be made of the information contained herein.

References

1. D. L. Ermak and J. A. McCammon, *Brownian dynamics with hydrodynamic interactions*, J. Chem. Phys., **69**, no. 4, 1352, 1978.
2. R. R. Gabdoulline and R. C. Wade, *Simulation of the diffusional association of Barnase and Barstar*, Biophys. J., **72**, no. 5, 1917–1929, May 1997.
3. R. R. Gabdoulline and R. C. Wade, *Brownian dynamics simulation of protein-protein diffusional encounter.*, Methods, **14**, no. 3, 329–41, Mar. 1998.
4. R. R. Gabdoulline and R. C. Wade, *On the contributions of diffusion and thermal activation to electron transfer between Phormidium laminosum plastocyanin and cytochrome f: Brownian dynamics simulations with explicit modeling of nonpolar desolvation interactions and electron transfer event*, J. Am. Chem. Soc., **131**, no. 26, 9230–8, July 2009.
5. D. Motiejunas, R. R. Gabdoulline, T. Wang, A. Feldman-Salit, T. Johann, P. J. Winn, and R. C. Wade, *Protein-protein docking by simulating the process of association subject to biochemical constraints.*, Proteins, **71**, no. 4, 1955–69, June 2008.
6. D. B. Kokh, S. Corni, W. Peter, H. Martin, K.-E. Gottschalk, and R. C. Wade, *ProMetCS: An Atomistic Force Field for Modeling Protein-Metal Surface Interactions in a Continuum Aqueous Solvent*, J. Chem. Theor. Comput., **6**, 1753–1768, 2010.
7. P. Mereghetti, R. R. Gabdoulline, and R. C. Wade, *Brownian dynamics simulation of protein solutions: structural and dynamical properties.*, Biophys. J., **99**, no. 11, 3782–91, Dec. 2010.
8. P. Mereghetti and R. Wade, *Diffusion of hydrophobin proteins in solution and interactions with a graphite surface*, BMC Biophys., **4**, no. 1, 9, 2011.
9. J. Balbo, P. Mereghetti, D.-P. Herten, and R. C. Wade, *The Shape of Protein Crowders is a Major Determinant of Protein Diffusion*, Biophys. J., **104**, no. 7, 1576 – 1584, 2013.
10. P. Mereghetti and R. C. Wade, *Atomic detail brownian dynamics simulations of concentrated protein solutions with a mean field treatment of hydrodynamic interactions.*, J. Phys. Chem. B, **116**, no. 29, 8523–33, July 2012.
11. P. Ren, J. Chun, D. G. Thomas, M. J. Schnieders, M. Marucho, J. Zhang, and N. A. Baker, *Biomolecular electrostatics and solvation: a computational perspective*, Q. Rev. Biophys., **45**, 427–491, 11 2012.
12. B. Roux and T. Simonson, *Implicit solvent models*, Biophys. Chem., **78**, no. 1-2, 1–20, Apr. 1999.
13. S. A. Hassan and E. L. Mehler, *A critical analysis of continuum electrostatics: the screened Coulomb potential–implicit solvent model and the study of the alanine*

dipeptide and discrimination of misfolded structures of proteins., Proteins, **47**, no. 1, 45–61, Apr. 2002.
14. T. Simonson and D. Perahia, *Internal and interfacial dielectric properties of cytochrome c from molecular dynamics in aqueous solution*, Proc. Nat. Acad. Sci. USA, **92**, no. February, 1082–1086, 1995.
15. A. Warshel, P. K. Sharma, M. Kato, and W. W Parson, *Modeling electrostatic effects in proteins.*, Biochim. Biophys. Acta, **1764**, no. 11, 1647–1676, Nov. 2006.
16. R. R. Gabdoulline and R. C. Wade, *Effective Charges for Macromolecules in Solvent*, J. Phys. Chem., **100**, no. 9, 3868–3878, Jan. 1996.
17. S. R. McGuffee and A. H. Elcock, *Atomically detailed simulations of concentrated protein solutions: the effects of salt, pH, point mutations, and protein concentration in simulations of 1000-molecule systems.*, J. Am. Chem. Soc., **128**, no. 37, 12098–12110, Sept. 2006.
18. D. A. Beard and T. Schlick, *Modeling salt-mediated electrostatics of macromolecules: The discrete surface charge optimization algorithm and its application to the nucleosome*, Biopolymers, **58**, 106–115, 2001.
19. D. A. Beard and T. Schlick, *Computational modeling predicts the structure and dynamics of chromatin fiber*, Structure, **9**, no. 01, 105–114, 2001.
20. G. A. Huber and J. A. McCammon, *Browndye: A Software Package for Brownian Dynamics.*, Comp. Phys. Comm., **181**, no. 11, 1896–1905, Nov. 2010.
21. M. Długosz, P. Zieliński, and J. Trylska, *Brownian dynamics simulations on CPU and GPU with BD_BOX*, J. Comput. Chem., **32**, no. 12, 2734–2744, 2011.
22. J. G. De La Torre, José G. Hernández C., A. Ortega, R. R. Schmidt, M. X. Fernandes, H. E. Pérez Sánchez, and R. Pamies, *SIMUFLEX: Algorithms and Tools for Simulation of the Conformation and Dynamics of Flexible Molecules and Nanoparticles in Dilute Solution*, J. Chem. Theor. Comput., **5**, no. 10, 2606–2618, Oct. 2009.
23. T. Geyer, *Many-particle Brownian and Langevin Dynamics Simulations with the Brownmove package*, BMC Biophys., **4**, no. 1, 7, 2011.
24. P. J. Goodford, *A computational procedure for determining energetically favorable binding sites on biologically important macromolecules*, J. Med. Chem., **28**, no. 7, 849–857, July 1985.
25. A. Onufriev, D. Bashford, and D. A. Case, *Modification of the Generalized Born Model Suitable for Macromolecules*, J. Phys. Chem. B, **104**, no. 15, 3712–3720, Apr. 2000.
26. A. Onufriev, D. A. Case, and D. Bashford, *Effective Born radii in the generalized Born approximation: the importance of being perfect.*, J. Comput. Chem., **23**, no. 14, 1297–304, Nov. 2002.
27. F. Floris and J. Tomasi, *Evaluation of the dispersion contribution to the solvation energy. A simple computational model in the continuum approximation*, J. Comput. Chem., **10**, no. 5, 616–627, July 1989.
28. F. M. Floris, J. Tomasi, and J. L. Pascual Ahuir, *Dispersion and repulsion contributions to the solvation energy: Refinements to a simple computational model in the continuum approximation*, J. Comput. Chem., **12**, no. 7, 784–791, Sept. 1991.
29. R. M. Levy, L. Y. Zhang, E. Gallicchio, and A. K. Felts, *On the nonpolar hydration free energy of proteins: surface area and continuum solvent models for the solute-solvent interaction energy.*, J. Am. Chem. Soc., **125**, no. 31, 9523–9530, Aug. 2003.
30. C. Tan, Y.-H. Tan, and R. Luo, *Implicit Nonpolar Solvent Models*, J. Phys. Chem. B,

111, no. 42, 12263–12274, 2007.
31. E. Gallicchio, L. Y. Zhang, and R. M. Levy, *The SGB/NP hydration free energy model based on the surface generalized born solvent reaction field and novel nonpolar hydration free energy estimators.*, J. Comput. Chem., **23**, no. 5, 517–529, Apr. 2002.
32. J. A. Wagoner and N. A. Baker, *Assessing implicit models for nonpolar mean solvation forces: the importance of dispersion and volume terms.*, Proc. Nat. Acad. Sci. USA, **103**, no. 22, 8331–8336, May 2006.
33. J. Dzubiella, J. M. J. Swanson, and J. A. McCammon, *Coupling nonpolar and polar solvation free energies in implicit solvent models.*, J. Chem. Phys., **124**, no. 8, 084905, Feb. 2006.
34. J. Dzubiella, J. Swanson, and J. McCammon, *Coupling Hydrophobicity, Dispersion, and Electrostatics in Continuum Solvent Models*, Phys. Rev. Lett., **96**, no. 8, 087802, Mar. 2006.
35. S. Zhou, L. Cheng, J. Dzubiella, B. Li, and J. A. McCammon, *Variational Implicit Solvation with Poisson-Boltzmann Theory.*, J. Chem. Theor. and Comput., **10**, no. 4, 1454–1467, Apr. 2014.
36. P. Mereghetti, M. Martinez, and R. C. Wade, *Long range Debye-Hückel correction for computation of grid-based electrostatic forces between biomacromolecules.*, BMC Biophys., **7**, 4, Jan. 2014.
37. A. H. Elcock, R. R. Gabdoulline, R. C. Wade, and J. A. McCammon, *Computer simulation of protein-protein association kinetics: acetylcholinesterase-fasciculin*, J. Mol. Biol., **291**, no. 1, 149 – 162, 1999.
38. M. Feig and C. L. Brooks, *Recent advances in the development and application of implicit solvent models in biomolecule simulations.*, Current opinion in structural biology, **14**, no. 2, 217–224, Apr. 2004.
39. A. Morreale, R. Gil-Redondo, and A. R. Ortiz, *A new implicit solvent model for protein-ligand docking.*, Proteins, **67**, no. 3, 606–16, May 2007.
40. C N Schutz and A. Warshel, *What are the dielectric "constants" of proteins and how to validate electrostatic models?*, Proteins, **44**, no. 4, 400–17, Sept. 2001.
41. N. Froloff, A. Windemuth, and B. Honig, *On the calculation of binding free energies using continuum methods: application to MHC class I protein-peptide interactions.*, Protein science : a publication of the Protein Society, **6**, no. 6, 1293–1301, June 1997.
42. J. Antosiewicz, J. A. McCammon, and M. K. Gilson, *The determinants of pKas in proteins.*, Biochemistry, **35**, no. 24, 7819–33, June 1996.
43. E. Demchuk and R. C. Wade, *Improving the Continuum Dielectric Approach to Calculating pK_as of Ionizable Groups in Proteins*, J. Phys. Chem., **100**, no. 43, 17373–17387, Jan. 1996.
44. T. Hou, W. Zhang, Q. Huang, and X. Xu, *An extended aqueous solvation model based on atom-weighted solvent accessible surface areas: SAWSA v2.0 model.*, J. Mod. Model., **11**, no. 1, 26–40, Feb. 2005.
45. T. Wang, S. Tomic, R. R. Gabdoulline, and R. C. Wade, *How optimal are the binding energetics of barnase and barstar?*, Biophys. J., **87**, no. 3, 1618–30, Sept. 2004.
46. F. Dong, M. Vijayakumar, and H.-X. Zhou, *Comparison of calculation and experiment implicates significant electrostatic contributions to the binding stability of barnase and barstar.*, Biophys. J., **85**, no. 1, 49–60, July 2003.
47. N. A. Baker, *Improving implicit solvent simulations: a Poisson-centric view.*, Current

opinion in structural biology, **15**, no. 2, 137–143, Apr. 2005.
48. W. Im, J. Chen, and C. L. Brooks, *Peptide and protein folding and conformational equilibria: theoretical treatment of electrostatics and hydrogen bonding with implicit solvent models.*, Adv. Protein Chem., **72**, no. 05, 173–98, Jan. 2005.
49. A. Onufriev, D. Bashford, and D. A. Case, *Exploring protein native states and large-scale conformational changes with a modified generalized born model.*, Proteins, **55**, no. 2, 383–94, May 2004.
50. D. Sitkoff, K. A. Sharp, and B. Honig, *Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models*, J. Phys Chem., **98**, no. 7, 1978–1988, Feb. 1994.
51. M. Ahmad, W. Gu, T. Geyer, and V. Helms, *Adhesive water networks facilitate binding of protein interfaces.*, Nat. Comm., **2**, 261, Jan. 2011.
52. S. A.. Hassan, F. Guarnieri, and E. L. Mehler, *A General Treatment of Solvent Effects Based on Screened Coulomb Potentials*, J. Phys. Chem. B, **104**, no. 27, 6478–6489, July 2000.
53. S. Lee, K.-H. Cho, Y.-M. Kang, H. A. Scheraga, and K. T. No, *A generalized G-SFED continuum solvation free energy calculation model.*, Proc. Nat. Acad. Sci. USA, **110**, no. 8, E662–7, Feb. 2013.
54. A. Onufriev, "Implicit solvent models in molecular dynamics simulations: A brief overview", in: Ann. Rep. Comp. Chem., vol. 4, pp. 125–137. 2008.
55. A. Rubinstein and S. Sherman, *Evaluation of the influence of the internal aqueous solvent structure on electrostatic interactions at the protein-solvent interface by nonlocal continuum electrostatic approach.*, Biopolymers, **87**, no. 2-3, 149–64, 2007.
56. J. H. Lin, N. A. Baker, and J. A. McCammon, *Bridging implicit and explicit solvent approaches for membrane electrostatics*, Biophys. J., **83**, no. September, 1374–1379, 2002.
57. S. Edwards, B. Corry, S. Kuyucak, and S. Chung, *Continuum electrostatics fails to describe ion permeation in the gramicidin channel*, Biophys. J., **83**, no. September, 1348–1360, 2002.
58. M. S. Lee, F. R. Salsbury, and M. A. Olson, *An efficient hybrid explicit/implicit solvent method for biomolecular simulations.*, J. Comput. Chem., **25**, no. 16, 1967–78, Dec. 2004.
59. V. Lounnas, S. K. Lüdemann, and R. C. Wade, *Towards molecular dynamics simulation of large proteins with a hydration shell at constant pressure*, Biophysical Chemistry, **78**, no. 1-2, 157–182, Apr. 1999.
60. N. V. Prabhu, M. Panda, Q. Yang, and K. A. Sharp, *Explicit ion, implicit water solvation for molecular dynamics of nucleic acids and highly charged molecules.*, J. Comput. Chem., **29**, no. 7, 1113–30, May 2008.
61. S. O. Yesylevskyy, L. V. Schäfer, D. Sengupta, and S. J. Marrink, *Polarizable water model for the coarse-grained MARTINI force field.*, PLoS Comp. Biol., **6**, no. 6, e1000810, June 2010.
62. Z. Wu, Q. Cui, and A. Yethiraj, *A new coarse-grained model for water: the importance of electrostatic interactions.*, J. Phys. Chem. B, **114**, no. 32, 10524–9, Aug. 2010.
63. D. J. Bonthuis and R. R. Netz, *Beyond the Continuum: How Molecular Solvent Structure Affects Electrostatics and Hydrodynamics at Solid-Electrolyte Interface*, J. Phys. Chem. B, **117**, 11397–11413, 2013.
64. S. Biggs, P. Mulvaney, C. F. Zukoski, and F. Grieser, *Study of Anion Adsorption at the*

Gold-Aqueous Solution Interface by Atomic Force Microscopy, J. Am. Chem. Soc., **116**, 9150–9157, 1994.
65. M. Henderson, *The interaction of water with solid surfaces: fundamental aspects revisited*, Surf. Sci. Rep., **46**, 1–308, 2002.
66. F. Iori, R. Di Felice, E. Molinari, and S. Corni, *GolP: An Atomistic Force-Field to Describe the Interaction of Proteins With Au(111) Surfaces in Water*, J. Comput. Chem, **30**, 1465, 2009.
67. F. Iori and S. Corni, *Including image charge effects in the molecular dynamics simulations of molecules on metal surfaces*, J. Comput. Chem., **29**, 1656, 2008.
68. K. Kubiak-Ossowska and P. A. Mulheran, *Protein diffusion and long-term adsorption States at charged solid surfaces*, Langmuir, **28**, 15577–15585, 2012.
69. T. Frembgen-Kesner and A. H. Elcock, *Absolute Protein-Protein Association Rate Constants from Flexible, Coarse-Grained Brownian Dynamics Simulations: The Role of Intermolecular Hydrodynamic Interactions in Barnase-Barstar Association*, Biophys. J., **99**, no. 9, L75 – L77, 2010.
70. M. Cieplak and S. Niewieczerzał, *Hydrodynamic interactions in protein folding*, J. Chem. Phys., **130**, no. 12, –, 2009.
71. T. Frembgen-Kesner and A. H. Elcock, *Striking Effects of Hydrodynamic Interactions on the Simulated Diffusion and Folding of Proteins*, J. Chem. Theor. and Comput., **5**, no. 2, 242–256, 2009.
72. T. Ando and J Skolnick, *On the Importance of Hydrodynamic Interactions in Lipid Membrane Formation*, Biophys. J., **104**, no. 1, 96–105, 2013.
73. J. Rotne and S. Prager, *Variational Treatment of Hydrodynamic Interaction in Polymers*, J. Chem. Phys., **50**, no. 11, 4831–4837, 1969.
74. J. G. Kirkwood, *The statistical mechanical theory of irreversible processes in solutions of flexible macromolecules. Visco-elastic behavior*, Recl. Trav. Chim Pay-B., **68**, no. 7, 649–660, 1949.
75. H. Yamakawa, *Transport Properties of Polymer Chains in Dilute Solution: Hydrodynamic Interaction*, J. Chem. Phys., **53**, no. 1, 436–443, 1970.
76. J. G. De La Torre and V. A. Bloomfield, *Hydrodynamic properties of macromolecular complexes. I. Translation*, Biopolymers, **16**, no. 8, 1747–1763, 1977.
77. M. Fixman, *Construction of Langevin forces in the simulation of hydrodynamic interaction*, Macromolecules, **19**, no. 4, 1204–1207, 1986.
78. T. Geyer and U. Winter, *An $O(N^2)$ approximation for hydrodynamic interactions in Brownian dynamics simulations*, J. Chem. Phys., **130**, no. 11, 114905, 2009.
79. A. H. Elcock, *Molecule-Centered Method for Accelerating the Calculation of Hydrodynamic Interactions in Brownian Dynamics Simulations Containing Many Flexible Biomolecules*, J. Chem. Theor. Comput., **9**, no. 7, 3224–3239, 2013.
80. D. M. Heyes, *Mean-field-hydrodynamics Brownian dynamics simulations of viscosity and self-diffusion of near-hard-sphere colloidal liquids*, J. Phys.-Condes. Matter, **7**, no. 47, 8857–8865, NOV 20 1995.

Continuum Solvation Modeling of Solute-Solvent Interactions

Robert Franke

Evonik Industries AG, Paul-Baumann-Str. 1, 45772 Marl, Germany
and
Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany

E-mail: robert.franke@evonik.com

Among the methods that model solvation, the class of continuum solvation models is characterized by rough approximations. These, however, enable quantitative predictions of properties such as solvation energies. The lecture notes are intended as a short overview of polarizable continuum models. After outlining the theory, the notes review some numerical examples. An extension of a certain polarizable continuum model is briefly presented. This extension allows quantitative predictions of thermophysical properties of mixtures.

1 Introduction

1.1 Preliminary remarks

The time constraints upon this presentation necessitate a strict selection of content for this topic. Continuum solvation models (CSMs) have been an intensively researched topic for decades; in particular, this class of methods is marked by the fact that they can be used in a wide variety of ways in application calculations. The temptation to illustrate the importance of this area by means of bibliometric time series plots is here resisted. Let it be noted, however, that a keyword search for the most important acronyms in this area of research in the literature databases will generate thousands of references. Many publications in which CSM methods are used are often not found in such searches, because it is now virtually the standard to also use CSM methods in experimental publications that are "supplemented" in particular by quantum chemical calculations, when the subject of study is the liquid phase(s). This certainly points toward a certain degree of maturity of these methods, although a great number of publications continue to be published that deal purely with the development of the method. An example of this is the recently published and probably the first completely formulated theory and the associated implementation of a CSM method in the context of a four-component relativistic Hartree-Fock and Kohn-Sham theory[1]. The selection criteria were related on the one hand to the technical requirements of a presentation and on the other hand were based on one of the focal points of the presenter's work, using the most economical means of calculation to provide data on physicochemical properties, which, at least in their trend, were quantitatively reliable. For this reason, the focus is on quantum mechanical and here on ab initio or density-functional-theory-based methods from the class of polarizable continuum models (PCMs). In particular in the area of industrial chemistry, agrochemical and pharmaceutical research, but also in the area of chemical engineering, PCM methods are being used increasingly, with particular relevance being attached to a semi-empirical, statistical-thermodynamic extension of a PCM, the so-called conductor-like screening model of real solvents (COSMO-RS) method[2] (see also a recently

published review article on the peer-reviewed publications on computational chemistry by researchers engaged exclusively in industry or presented with their participation[3]). This will be outlined briefly at the end of the lecture. The elegant integral equation formalism developed in the context of PCM theories, for example, will not be handled. Time-dependent phenomena will also be left out. The generalized Born approaches or methods based on finite element and finite difference techniques will also not be dealt with. The use of CSMs in molecular dynamics and Monte Carlo calculations with classical interatomic potentials or in ab initio molecular dynamics methods will also not be discussed. Technical aspects of concrete implementations will, for the most part, be left out. The availability of the individual methods in freely available or commercially obtainable program packages will also not be discussed. However, let it be noted that every popular quantum chemical program system contains at least one variant of the PCM methods dealt with here.

The literature to the year 2004 is very well reviewed through multiple review articles. Deserving of special mention are the publications by Tomasi and Persico[4], Cramer und Truhlar[5] as well as Tomasi, Mennucci, and Cammi[6]. The literature of the past ten years is less well developed. In addition to some reviews on special topics or technical aspects, some of which will be mentioned later, let us also note a collection from the year 2007[7] and a current review on PCM for multiconfigurational self-consistent field theory[8] .

1.2 Solvation free energy in continuum models

It is assumed that a solution is represented by a collection of molecules consisting of solute molecules M and solvent molecules S and in which only noncovalent molecular interactions take place, in particular no bond cleavage or bond formation. Furthermore a state of infinite dilution is assumed. For the electronic ground state and with validity of the Born-Openheimer approximation, the interactions in the system are described by the Hamilton operator

$$\hat{H}\left(\vec{R}_M, \vec{R}_S\right) = \hat{H}^M\left(\vec{R}_M\right) + \hat{H}^S\left(\vec{R}_S\right) + \hat{H}^{SS}\left(\vec{R}_S\right) + \hat{H}^{MS}\left(\vec{R}_M, \vec{R}_S\right) \quad (1)$$

where the coordinates $\vec{R}$ describe nuclei and electrons. $\hat{H}^{SS}$ describes the interaction between solvent molecules and $\hat{H}^{MS}$ between the solute molecule M and the solvent molecules. In continuum models a drastic approximation is now introduced for this description, which itself is also naturally a rough approximation, by describing the interactions in the collection of molecules using an effective Hamilton operator, which depends only on the coordinates of the one solute molecule $\vec{R}_M$

$$\hat{\tilde{H}}\left(\vec{R}_M\right) = \hat{H}^M\left(\vec{R}_M\right) + \hat{\tilde{H}}^{MS}\left(\vec{R}_M\right) \quad (2)$$

In a certain sense, this reduction of the coordinates represents the essential difference between a discrete description and one using continuum models. $\hat{\tilde{H}}^{MS}$ is a sum of operators that represent specific interactions. As a rule, four of these terms are taken into consideration, each representing a contribution to solvation free energy. In addition, there is also a basic discussion concerning inclusion of the contribution of the thermal motion of the molecular framework, ΔG_{tm}. The solvation free energy in continuum solvation models is now defined as follows:

$$\Delta G_{\mathrm{sol}} = G_{\mathrm{cav}} + \Delta G_{\mathrm{el}} + G_{\mathrm{rep}} + G_{\mathrm{dis}} + \Delta G_{\mathrm{tm}} + P\Delta V \quad (3)$$

This energy relates to a system in which the solute molecule M is in the phase of an ideal gas, and the solvent is in equilibrium as a pure liquid. In (3), G_{cav} refers to the so-called cavitation free energy. In the pure, undisturbed phase of the solvent, it must be used to create a space with the exact form that enables the solute molecule M to find room within it. For a quantitative description, one needs data that describe the three-dimensional shape and the filling of the space of molecule M. A method based on hard sphere theory is very often used for calculation of this contribution[8]. No further details will be discussed here. Alternative methods of calculation can be taken from[6]. ΔG_{el} is the electrostatic contribution to ΔG_{sol} and the central quantity in continuum solvation models. It goes back to changes of nuclear coordinates and electron distributions as a consequence of the electrostatic interactions between the molecule M and the solvent surrounding it, which is described as a structure-less, dielectric medium. In the following, we deal predominantly with the efficient and precise calculation of these quantities. G_{rep} and G_{dis} describe van der Waals contributions to interaction energy. These contributions are comparatively small. For details on estimation, please refer to[6]. As a rule the last two contributions, ΔG_{tm} and $P\Delta V$, are neglected[6].

2 Continuum Solvation Models

Very generally, an initial definition of continuum solvation models can be based on the principle of a qualitatively different treatment of solute and solvent. In the broadest sense, atomistic models (and this note focuses on quantum mechanical methods) are used to describe the solute. The solvent is described as a polarizable continuum. The continuum is homogeneous and isotropic and is in a state of equilibrium at a given temperature and given pressure. Interactions are limited to those of an electrostatic nature. This approach goes back to a classic work[9], which was published at a time at which it was still unclear whether there was such a thing as charged particles in solutions.

2.1 The Born model

In 1920, Max Born published a paper on the calculation of the free solvent energy of ions (in water), which was based on a very simplified description of the interaction between an ion dissolved in the liquid, which was viewed as a charged sphere with a given radius, and the solvent[9]. It will be presented briefly here, because it is very well suited for explaining the physical bases of CSMs. The fundamental assumptions are illustrated in Figure 1.

The discrete model, in which the solvent molecules (water) are viewed atomistically, is very greatly simplified in the Born model. The water molecules are replaced by a continuous polarizable medium, which extends to infinity. The polarizability is described by the (static) dielectric constant (relative permittivity) ε_r. In a vacuum it is equal to 1; for water at 25° C and 1 bar it is 78.5, and for typical nonpolar solvents it has a value of about 2. The permittivity is $\varepsilon = \varepsilon_r \cdot \varepsilon_0$, with the permittivity of the vacuum being ε_0. For quantitative description, the electrostatic energy in dielectrics is needed. Based on textbooks on classic electrodynamics (see e.g.[10–13]), certain relationships will be summarized in the following, which point in particular toward the underlying approximations used for CSMs. We consider the energy stored in an electrostatic system that contains a dielectric and proceed from the assumption that we are adding a free charge portion by portion. Because the free

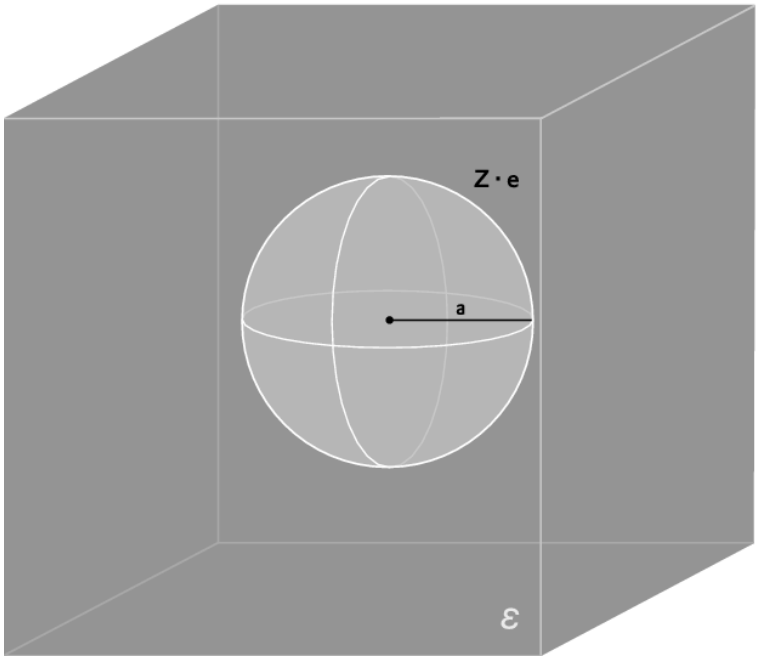


Figure 1: Sketch of the Born model.

charge distribution ρ is increased by an incremental amount $\delta\rho$, the polarization changes. In turn this changes the charge density of the bound charges in the dielectric and on its surface. The field created by the polarization of the medium corresponds to this charge density. However, the only thing that is of interest initially is the work done at the free charge:

$$\delta W = \int d\tau \, (\delta\rho) \, V \tag{4}$$

Here, V refers to the electrostatic potential and $d\tau$ to a volume element. In accordance with Gauss's law, formulated for the dielectric displacement $\vec{D}$ the following applies: $\nabla\vec{D} = \rho$ and thus: $\delta\rho = \nabla\left(\delta\vec{D}\right)$ so that the following applies for (4):

$$\delta W = \int d\tau \left[\nabla\left(\delta\vec{D}\right)\right] V \tag{5}$$

With $\nabla\left[\left(\delta\vec{D}\right)V\right] = \left[\nabla\left(\delta\vec{D}\right)\right]V + \delta\vec{D}\,(\nabla V)$ and $\vec{E} = -\nabla V$ it is possible to write for the integral (5):

$$\delta W = \int d\tau \, \nabla\left[\left(\delta\vec{D}\right)V\right] + \int d\tau \, \left(\delta\vec{D}\right)\vec{E} = 0 + \int d\tau \, \left(\delta\vec{D}\right)\vec{E} \tag{6}$$

Here, the first integral after the first equal sign in (6) is transformed into a surface integral using Gauss's law, which for the case $V\,(r)$ disappears at least like $1/r$. Equation (6)

applies generally and can be applied to any materials. If one assumes an isotropic, homogeneous (ε is not a function of the position), linear dielectric, then for the dielectric displacement and the electric field, the following relationship applies: $\vec{D} = \varepsilon\vec{E}$, so that the integrand (6) can be transformed into: $\left(\delta\vec{D}\right)\vec{E} = \varepsilon\Big/2 \cdot \delta\left(\vec{E}\vec{E}\right) = 1\Big/2 \cdot \delta\left(\vec{D}\vec{E}\right)$ and, for the entire work, results in:

$$W = \frac{1}{2}\int d\tau\, \vec{D}\vec{E} \tag{7}$$

Here a brief excursion about the approximate character of the description may be interjected. The assumption that the dielectric shows linear response and that the relationship $\vec{D} = \varepsilon\vec{E}$ applies is an approximation, which can fail for large $\vec{E}$. For example, for small, multiply charged ions, this can cause errors[14]. Depending on the material, the proportionality can cause large errors even for small fields. This relationship may also no longer be correct for fields that change over time. Furthermore, the assumed isotropy must be questioned, for example in the case of polar solvents. This approximation must thus always be kept in mind when using CSMs. Proceeding from equation (4), we now examine the Born model more closely. As shown in Figure 1, the ion is described as a point charge ze with the valency of the ion z and the elementary charge e, as well as the radius a. The electrostatic potential is then

$$V(r) = \frac{ze}{4\pi\varepsilon}\frac{1}{r} = \frac{ze}{4\pi\varepsilon_r\varepsilon_0}\frac{1}{r} \tag{8}$$

For the electric field, the following applies.

$$\vec{E} = -\nabla V = \frac{ze}{4\pi\varepsilon_r\varepsilon_0}\frac{\vec{r}}{r^3} \tag{9}$$

The electrostatic energy of the system is thus:

$$W = \frac{1}{2}\int d\tau\, \vec{D}\vec{E} = \frac{4\pi}{2}\int_a^\infty dr\, r^2 \frac{z^2e^2}{16\pi^2\varepsilon_r\varepsilon_0}\frac{r^2}{r^6} = \frac{z^2e^2}{8\pi\varepsilon_r\varepsilon_0}\left[-\frac{1}{r}\right]_a^\infty = \frac{z^2e^2}{8\pi\varepsilon_r\varepsilon_0}\frac{1}{a} \tag{10}$$

In a vacuum, in equation (10) $\varepsilon_r = 1$ must be set. Now we calculate the energy of the transfer of an ion from the vacuum into a medium, characterized by its dielectric constant, meaning a free energy of the solvation in the standard state of ideally diluted solutions at a given temperature and pressure. Henry's Law applies (see e. g.[15]). Only the interaction between ion and solvent is described. Interactions among the solvent molecules and those among the ions are not taken into consideration; nor are entropic contributions, e. g. those caused by dilution effects. For the solvation energy according to Born, this means:

$$\Delta G^0_{\mathrm{Born}} = W(\varepsilon_r) - W(\varepsilon_r = 1) = \frac{z^2e^2}{8\pi\varepsilon_0}\left(\frac{1}{\varepsilon_r} - 1\right)\frac{1}{a} = -\frac{z^2e^2}{8\pi\varepsilon_0}\left(1 - \frac{1}{\varepsilon_r}\right)\frac{1}{a} \tag{11}$$

The Born model shows clear quantitative weaknesses, which should not be surprising, considering the rough approximations. For example, the inverse dependency of the ion radius upon the solvation energy cannot be confirmed experimentally. Empirical correction of the ion radii (addition of a constant quantity to the ion radius, which differs for cation and anion) can save the relationship. The solvation energies calculated with the Born model continue to be generally too high. This is a consequence of the approximation employing

a dielectric constant which is independent of position. This can also be corrected through ad hoc extensions to the model. Despite substantial quantitative weaknesses, particularly the conceptional importance of this model can hardly be overestimated. The statements in the last two sentences in Max Born's publication: "Nun ist aber sowohl die Berechnung der Hydratationswärmen viel zu ungenau, als auch die Definition der Ionen- und Atomradien viel zu unbestimmt, als daß man diesem quantitativen Ergebnisse irgendwelches Vertrauen entgegenbringen dürfte. Ich glaubte aber doch, die einfache Rechnung mitteilen zu müssen, weil sie auf jeden Fall geeignet ist, die elektrostatische Auffassung des Hydratationsvorganges zu stützen." (The calculation of the heat of hydration is much too inexact and the definition of the ion and atomic radii much to imprecise to permit any confidence in these quantitative results. Nevertheless, I believe that I have to report the simple calculation, because it is suited in any case to support the electrostatic understanding of the process of hydration.) now seem very modest in historical perspective.

2.2 The Onsager model

In 1936 Lars Onsager published a very influential work[16], in which he critically examined and expanded the Debye equation[17], which links the permittivity with the molecular properties of dipole moment and polarizability. The important concept of a "reaction field" and the term itself go back to this work. The Onsager model is contained in many quantum chemical program packages and is used for a wide variety of applications. For an attempt to put it into historical context, please refer to[18]. In this model, the solute is viewed as a point dipole $\vec{\mu}$ which can be polarized and is located in a spherical cavity with a radius of a. The cavity is embedded in a dielectric continuum, which is characterized by the dielectric constant of the solvent (see Figure 2). Since many questions in CSMs can be discussed well using this model problem, we present a detailed discussion of it here.

The basic problem of electrostatics is the solution of the Poisson equation, a linear, inhomogeneous, partial differential equation.

$$\Delta V\left(\vec{r}\right) = -\frac{1}{\varepsilon_0} \cdot \rho\left(\vec{r}\right) \tag{12}$$

The solution of equation (12) for a given charge density $\rho\left(\vec{r}\right)$ is one of the main problems in working with CSMs. In principle, an accurate solution of the Poisson equation provides the exact answer to the electrostatic problem, which is given by a concrete CSM. Equation (12) basically provides many solutions, which depend upon the boundary conditions. Since in the Onsager model it is the solutions inside and outside the cavity (but not on the surface) that are of interest, meaning the electrostatic potential $V\left(\vec{r}\right)$ in spatial areas in which the charge density is $\rho\left(\vec{r}\right) = 0$, a special case of the Poisson equation must be solved, the Laplace equation.

$$\Delta V\left(\vec{r}\right) = 0 \tag{13}$$

Spherical coordinates are used. The dipole $\vec{\mu}$ points in the +z direction; so that: $\vec{\mu} = \mu\hat{z}$. Since the potential on the surface of the sphere must be constant, the following applies:

$$V_{in}\left(a\right) = V_{out}\left(a\right) \tag{14}$$

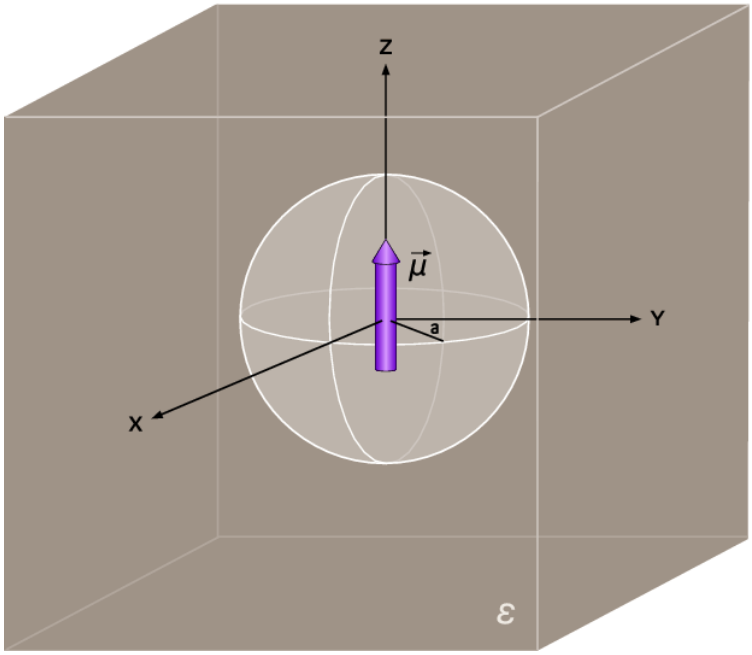


Figure 2: Sketch of the Onsager model.

There are no free charges on the surface, so that the following applies for a linear dielectric

$$\varepsilon_0 \frac{\partial V_{in}(r)}{\partial r}\bigg|_{r=a} = \varepsilon \frac{\partial V_{out}(r)}{\partial r}\bigg|_{r=a} \tag{15}$$

Since there is no external field, the following applies for the potential:

$$\lim_{r\to\infty} V(r) = 0 \tag{16}$$

For $r \to 0$ the potential should behave like that of an ideal dipole, which means that:

$$\lim_{r\to 0} V(r) = \frac{\mu \cdot \cos\theta}{4\pi\varepsilon_0 r^2} \tag{17}$$

In the case of azimuthal symmetry, equation (13) has the general solution[12,13]

$$V(r,\theta) = \sum_{l=0}^{\infty} \left(A_l r^l + \frac{B_l}{r^{l+1}} \right) P_l(\cos\theta) \tag{18}$$

Where P_l designates the Legendre polynomials. Under the conditions (16) and (17), the following solutions apply for the electrostatic potential inside and outside the cavity:

$$V_{in}(r,\theta) = \frac{\mu \cdot \cos\theta}{4\pi\varepsilon_0 r^2} + \sum_{l=0}^{\infty} A_l r^l P_l(\cos\theta) \tag{19}$$

$$V_{out}(r,\theta) = \sum_{l=0}^{\infty} \frac{B_l}{r^{l+1}} P_l(\cos\theta) \tag{20}$$

Inserting (19) and (20) in (14) and equating the coefficients results in:

$$B_1 = \frac{\mu}{4\pi\varepsilon_0} + A_1 a^3;\ B_l = a^{2l+1} A_l \ \text{ for } (l \neq 1) \tag{21}$$

(15) results in the relations:

$$-\frac{\mu}{2\pi a^3} + \varepsilon_0 A_1 = -2\varepsilon \frac{B_1}{a^3} \tag{22}$$

$$\varepsilon_0 l A_l a^{l-1} = -\varepsilon\,(l+1)\,\frac{B_l}{a^{l+2}} \text{ for } (l \neq 1) \tag{23}$$

Inserting B_l from (21) into (23) results in the equation

$$\varepsilon_0 l A_l a^{2l+1} = -\varepsilon\,(l+1)\,A_l a^{2l+1} \text{ for } (l \neq 1) \tag{24}$$

This can only be satisfied, if $A_l = 0$ applies for all $(l \neq 1)$. Thus $A_l = B_l = 0$ also applies for all $(l \neq 1)$. For the electrostatic potentials (equations (19) and (20)) this means:

$$V_{in}(r,\theta) = \frac{\mu \cdot \cos\theta}{4\pi\varepsilon_0 r^2} + A_1 \cdot r\cos\theta \tag{25}$$

$$V_{out}(r,\theta) = B_1 \frac{\cos\theta}{r^2} \tag{26}$$

From equations (21) and (22), it is possible to determine the coefficients A_1 and B_1:

$$A_1 = \frac{\mu}{2\pi\varepsilon_0 a^3} \cdot \frac{\varepsilon_0 - \varepsilon}{2\varepsilon + \varepsilon_0} = \frac{\mu}{2\pi\varepsilon_0 a^3} \cdot \frac{1-\varepsilon_r}{2\varepsilon_r + 1} \tag{27}$$

$$B_1 = \frac{\mu}{4\pi\varepsilon_0} \cdot \frac{3\varepsilon_0}{2\varepsilon + \varepsilon_0} = \frac{\mu}{4\pi\varepsilon_0} \cdot \frac{3}{2\varepsilon_r + 1} \tag{28}$$

The electrostatic potentials inside and outside the sphere are finally calculated from (25) and (26) as well as (20) as:

$$V_{in}(r,\theta) = \frac{\mu \cdot \cos\theta}{4\pi\varepsilon_0 r^2} + \frac{\mu}{2\pi\varepsilon_0 a^3} \cdot \frac{1-\varepsilon_r}{2\varepsilon_r + 1} \cdot r\cos\theta = \frac{\mu \cdot \cos\theta}{4\pi\varepsilon_0 r^2} - R \cdot r\cos\theta \tag{29}$$

where

$$R = \frac{\mu}{2\pi\varepsilon_0 a^3} \cdot \frac{\varepsilon_r - 1}{2\varepsilon_r + 1} \tag{30}$$

as well as the following definition:

$$R = g \cdot \mu; \quad g = \frac{1}{2\pi\varepsilon_0 a^3} \cdot \frac{\varepsilon_r - 1}{2\varepsilon_r + 1} \tag{31}$$

and

$$V_{out}(r,\theta) = \frac{\mu}{4\pi\varepsilon_0} \cdot \frac{3}{2\varepsilon_r + 1} \cdot \frac{\cos\theta}{r^2} \tag{32}$$

The coefficient B_1 multiplied by ε_r is designated as the external moment of the dipole. This is based on the formal consideration of the field in the dielectric as a virtual dipole in the center of the cavity[19]. The coefficient R describes the electric field that acts upon the dipole. The field results from the electric displacements that are induced by the dipole. It is referred to as a reaction field. It acts in the spatial areas occupied by the solute. In Onsager's reaction field theory, the interaction between solute and solvent is thus described as follows: The solute (characterized by its dipole moment and its size, described by the sphere's radius a) induces a dipole in the surrounding dielectric medium, and the electric field generated in this way in turn acts upon the dipole of the solute (the reaction).

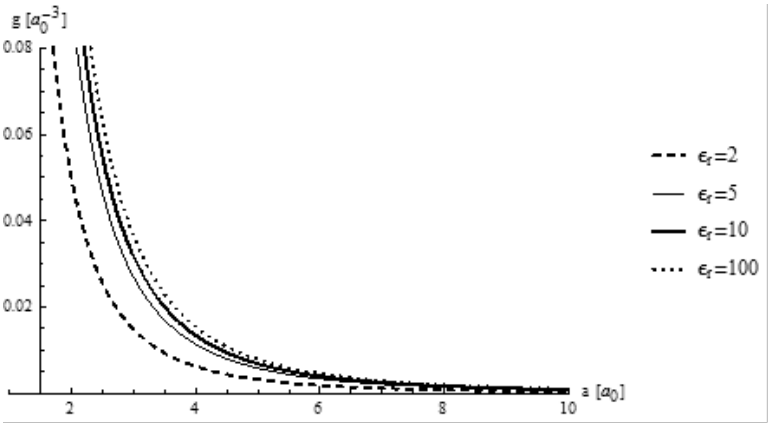


Figure 3: The factor g according to (31) as a function of the sphere's radius a and parametrically dependent upon ε_r in units a_0^3 with $a_0 = 0.529177 \cdot 10^{-10}\mathrm{m}$.

The codomain and shape of the proportionality factor g, which determines the strength of the reaction fields in the Onsager model, and which depends upon the sphere's radius a and the dielectric constant ε_r, can be taken from Figure 3. It is of note here that g depends strongly upon the radius of the cavity. In contrast, the dependency on the dielectric constant is of a subordinate nature. Of course, this behavior is also reflected in studies with the Onsager model. An example is a very impressive publication on trends in silicon antranes[20].

The electrostatic contribution to the solvation energy is calculated from the energy of a dipole in an electric field as

$$\Delta G^0_{\mathrm{Onsager}} = -\mu \cdot R + \frac{1}{2}\mu \cdot R = -\frac{1}{2}\mu \cdot R = -\mu \cdot \frac{\mu}{4\pi\varepsilon_0 a^3} \cdot \frac{\varepsilon_r - 1}{2\varepsilon_r + 1} \tag{33}$$

which takes into consideration that work must be done in order to polarize the dielectric. For a linear dielectic, this amounts to $1/2 \cdot \mu \cdot R$. The classic electrostatic theory of linear dielectrics leads very generally to the relationship between the free energy for the solute-solvent interaction $G_{\mathrm{solute-solvent}}$ and the intrasolvent interaction $G_{\mathrm{intrasolvent}}$[19]

$$\Delta G_{\mathrm{intrasolvent}} = -\frac{1}{2}\,\Delta G_{\mathrm{solute-solvent}} \tag{34}$$

After the solute has polarized the solvent and has thus decreased the solute-solvent energy by $\Delta G_{\mathrm{solute-solvent}}$, half this gain in free energy is used up by the work that must be done in order to polarize the solvent. The free energy of the solute rises.

Onsager's model can be implemented relatively simply in quantum chemical programs. This method is known as a "self-consistent reaction field" (SCRF)[21]. Here, the electrostatic interaction between solute and solvent is included as an additional term in the Hamilton operator of the isolated molecule. The reaction field operator has the following form:

$$\hat{H}_{RF} = \hat{H}_0 + \hat{H}_1 \tag{35}$$

where $\hat{H}_0$ designates the Hamilton operator in a vacuum, and the operator $\hat{H}_1$ (an one-electron operator) designates the link between the molecular dipole operator $\hat{\vec{\mu}}$ and the reaction field $\vec{R}$:

$$\hat{H}_1 = -\hat{\vec{\mu}}\vec{R} \tag{36}$$

Equation (36) describes the interaction of the solute with the reaction field. A self-consistent description of this interaction is possible by incorporating an additional term in the Fock operator. The Fock matrix then has the form:

$$F_{\alpha\beta} = F^0_{\alpha\beta} - \vec{R} \cdot \left\langle \varphi_\alpha \mid \hat{\vec{\mu}} \mid \varphi_\beta \right\rangle = F^0_{\alpha\beta} - g \cdot \vec{\mu} \cdot \left\langle \varphi_\alpha \mid \hat{\vec{\mu}} \mid \varphi_\beta \right\rangle \tag{37}$$

with the basis functions $\varphi_\alpha$and φ_β. The total energy is calculated as:

$$E_{\mathrm{SCRF}} = \left\langle \Psi \mid \hat{H}_0 \mid \Psi \right\rangle - \frac{1}{2}\vec{\mu} \cdot \vec{R} \tag{38}$$

where Ψ is the total wave function of the molecule, calculated self-consistently for $\vec{\mu}$ and $\vec{R}$ using the Fock operators as per (37). For molecules without a permanent dipole moment, E_{SCRF} is identical to the energy in the vacuum. The Hartree-Fock method modified according to equation (37) results in an increase in the dipole moment in increasingly polar solvents. The generic algorithm for execution of an SCRF calculation in the Hartree-Fock (HF) approximation or on the density functional (DFT) level looks as follows:

1. Specification of $\vec{R}$
2. Self-consistent Hartree-Fock or Kohn-Sham calculation and calculation of the dipole moment with applied reaction field $\vec{R}$ (standard calculation with Fock operator according to (37)).
3. Calculation of an improved reaction field $\vec{R} = g\vec{\mu}$
4. If no convergence: back to 2.

This method can also be applied very easily on the correlation level like second order Møller-Plesset (MP2), or coupled cluster (CC). Due to the high efficiency of this method, numerous applications can be found in the literature. We refer particularly to the review[4] .

To illustrate the numerical characteristics of this method, some results on the conformational equilibrium between trans- and gauche-1,2-dichloroethane are discussed in the following. The quantum chemical calculations for this are done with the 6-31+G* basis. The geometries are based on SCRF calculations done on the Hartree-Fock level with a 6-31G* basis set. In the gas phase it is found experimentally that the energy of the

gauche-form is about 5kJ/mol higher than the trans-form. In polar solvents this difference disappears almost completely. These experimental findings[22] are compared in Figure 4 to the calculated energy differences[23]. In the Onsager model, the only adaptable parameter is the radius of the cavity. In the calculations it is set to $a = 3.65\,\text{Å}$[23]. If one considers the vacuum, one sees that the energy difference is overestimated both on the Hartree-Fock and on the MP2 level, although MP2 shows somewhat better agreement with the experiment. On the SCRF level, it is seen that both methods correctly reflect the qualitative course of the dependencies of the solute under consideration and that, as in the vacuum, MP2 provides better agreement. However, unlike the vacuum, the energy difference is underestimated here.

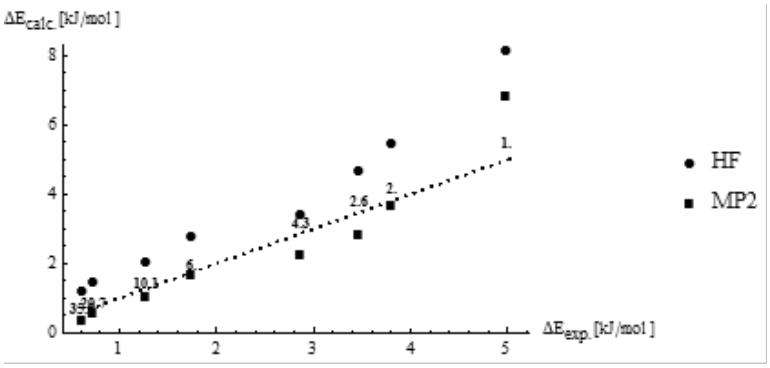


Figure 4: Calculated energy differences in different solvents of the conformers of the 1,2-dichloroethane $\Delta E = E_{\text{gauche}} - E_{\text{trans}}$ as a function of the experimental energy differences. The numerical values on the angle bisector correspond to the dielectric constant ε_r of the various solvents. They are (in descending order for the dielectric constant): acetonitrile, acetone, 1,2-dichloroethane, ethylacetate, diethylether, carbon disulfide, cyclohexane, vacuum. All numerical values as per[23].

In practical calculations with the Onsager model, the selection of the sphere diameter represents the greatest problem. There are numerous investigations on this topic. For example, possible specifications are based on the experimental density of the solute, the determination of the maximum length of the solute molecule, and subsequent addition of the van der Waal radii of the outermost atoms. As an example, we refer to a publication that uses a combination of the classic field-theoretical treatment of dielectrics and SCRF theory to derive a unique criterion for determination of the radius[24]. For the example presented here of 1,2-dichloroethane, the molar density provides a value of $a = 3.15\,\text{Å}$[23]. A consideration of the experimental values for the energy difference of the conformers shows that the value approaches zero as the polarity of the solvent increases[22]. Examining the energy difference in water as a solvent as a function of the cavity radius now leads to the approach shown in Figure 5[23]. Of note here is the strong dependence of the energy difference upon the selected radius. It decreases as the radius increases, whereas the dipole moment increases. This reflects the dependence of the factor g upon the radius a in accordance with equation (31) (see also Figure 3). From the behavior in Figure 5, it can now be concluded that values of $3.25\,\text{Å} < a < 3.75\,\text{Å}$ reflect the experimental values.

Finally, let us summarize the strengths and weaknesses of the Onsager model. The

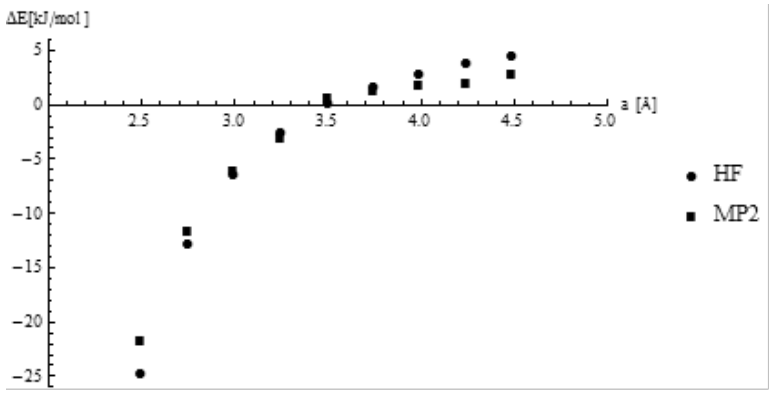


Figure 5: Calculated energy differences of the conformers of 1,2-dichloroethane $\Delta E = E_{\mathrm{gauche}} - E_{\mathrm{trans}}$ as a function of cavity radius a. The calculations are done for $\varepsilon_r = 78.5$.

great strength of the model is its conceptual simplicity and the characteristic of extreme ease of implementation into quantum chemical programs. However, the price for this is the rather unrealistic description. As a rule, the form of the cavity is described as a sphere. Onsager already mentioned this weakness in his own work and proposed expanding it toward an ellipsoid form of cavity, which was indeed examined later on. Access to the corresponding literature can be found, e. g. via[4]. However, this is still a very rough description of the as a rule irregular shape of the cavities of typical molecules. Furthermore, as a rule, the electrostatic potential of a solute is poorly approximated by a dipole moment. Higher multipoles play a not inconsiderable role (see e.g.[25]). The correct behavior of the model in the limiting case $\varepsilon_r \to \infty$ is also questionable[26].

2.3 The polarizable continuum model

A distinctly more realistic model is the polarizable continuum model (PCM). In this model, the solute molecule is placed in a cavity that is created by a set of spheres, whose central points each represent the position of the atom in the molecule. Thus the shape of the cavity has a form that approximates much more closely the individual characteristics of a molecule than, for example, assuming a sphere or an ellipsoid. The electrostatic interaction of the solute with the medium that surrounds it is now modeled in comparatively more detail. The algorithm used for this purpose is based on the idea of dividing the cavity into small parts, so-called tesserae (Latin for quadrangles; refers to the tiles from which mosaics are made). Each of these tesserae now has a specific "apparent surface charge" (ASC), which describes the solute-solvent interaction and also the interaction between the tesserae themselves. The surface charges contain the complete electrostatic solvation effects. From them are derived the energy of interaction and all other properties calculated with the PCM. The numerical values of the apparent surface charges are obtained by solving the Poisson equation (12). The first publication that describes such an approach for ab initio methods is[27]. The various methods of PCM differ in the ways and means of solving this problem. This also includes the concrete method for generation of the tesserae. In analogy to equation (35), the effective Hamilton operator, while assuming the validity of

the Born-Oppenheimer approximation, describes the molecule M in a solvent

$$\hat{H}_M = \hat{H}_0 + \hat{V}(\Psi) \tag{39}$$

Here, $\hat{H}_0$ describes the electronic Hamilton operator of the molecule in a vacuum, and $\hat{V}(\Psi)$ is an operator that describes the solute-solvent interaction. It depends on the electronic charge density $\rho_M^e(\vec{r})$ of the wave function of the solute. In addition, the dielectric, which characterizes the solvent, is also polarized by the nuclear charges of the solute, and thus in total by

$$\rho_M = \rho_M^e + \rho_M^n \tag{40}$$

Analogous to the quantum mechanical implementation of the Onsager model, two linked problems need to be solved: the problem of classical electrostatics of determining the solvent reaction potential induced by ρ_M and the quantum mechanical problem of calculating the wave function Ψ of the solute, defined by (39).

2.3.1 Solution of the Poisson equation

The potential of the solute-solvent interaction in equation (39) is the sum of the potential V_M that is created by the electronic and nuclear charge distribution of the solute and the reaction field potential V_R created by the polarization of the dielectric medium.

$$V(\vec{r}) = V_M(\vec{r}) + V_R(\vec{r}) \tag{41}$$

The electrostatic problem to be solved can now be formulated generally as follows.

$$-\nabla(\epsilon_r(\vec{r})\nabla V(\vec{r})) = 4\pi\rho_M(\vec{r}) \tag{42}$$

where $\lim_{r\to\infty} V(r) = 0$. The potential of the molecule, which is created by ρ_M in a vacuum, is:

$$V_M(\vec{r}) = \int_{\mathbb{R}^3} d\vec{\tilde{r}} \frac{\rho_M\left(\vec{\tilde{r}}\right)}{|\vec{r} - \vec{\tilde{r}}|} \tag{43}$$

For the set of all points of the three-dimensional space that is defined by the cavity in the dielectric in which the solute molecules are located (see Figure 6), the designation C is introduced.

Now, $\epsilon_r(\vec{r}) = 1$ applies in C, and $\epsilon_r(\vec{r}) = \epsilon_r$ applies outside C. A case distinction thus follows for (42)

$$\begin{array}{cc} -\Delta V(\vec{r}) = 4\pi\rho_M(\vec{r}) & \text{in C}, \vec{r} \in C \\ -\Delta V(\vec{r}) = 0 & \text{outside C}, \vec{r} \notin C \\ V(\vec{r}) \longrightarrow 0 & \text{at infinity} \end{array} \tag{44}$$

The same applies analogously for $V_M(\vec{r})$

$$\begin{array}{cc} -\Delta V_M(\vec{r}) = 4\pi\rho_M(\vec{r}) & \text{in C}, \vec{r} \in C \\ -\Delta V_M(\vec{r}) = 0 & \text{outside C}, \vec{r} \notin C \\ V_M(\vec{r}) \longrightarrow 0 & \text{at infinity} \end{array} \tag{45}$$

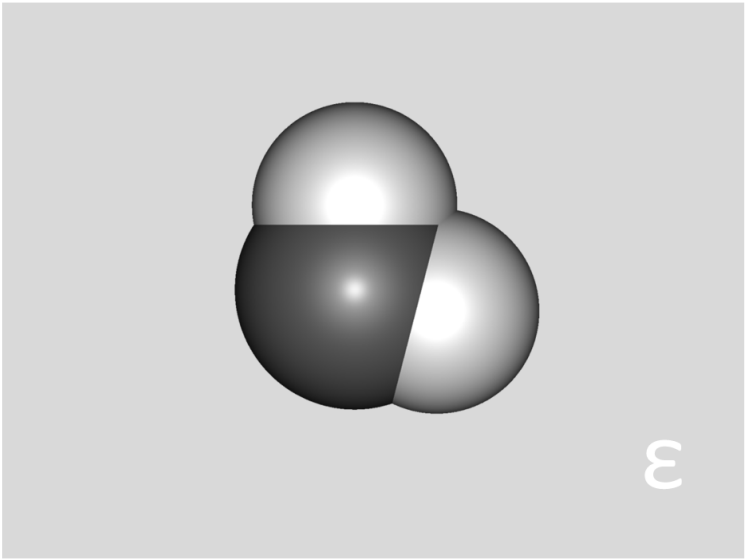


Figure 6: Sketch of the PCM for water.

Thus the following applies for $V_R(\vec{r})$ based on (41):

$$\begin{array}{rll} -\Delta V_R(\vec{r}) &= 0 & \text{in C}, \vec{r} \in C \\ -\Delta V_R(\vec{r}) &= 0 & \text{outside C}, \vec{r} \notin C \\ V_R(\vec{r}) &\longrightarrow 0 & \text{at infinity} \end{array} \tag{46}$$

Based on the characteristics of ρ_M (sum of point charges within C added to a regular function, the electronic density), it follows that $V_R(\vec{r})$ is continuous on the surface of C. This surface is designated as Γ. It is thus now possible to represent the reaction field potential in a compact way, which is of great practical significance. The following integral representation of the potential is the basic of the ASC approach to solving the Poisson equation

$$V_R(\vec{r}) = \int_\Gamma d\vec{s}\, \frac{\sigma(\vec{s})}{|\vec{r} - \vec{s}|} \qquad \text{for} \quad \vec{r} \in \mathbb{R}^3, \quad \vec{s} \in \Gamma \tag{47}$$

Here, $\sigma(\vec{s})$ is a distribution of charges on the surface Γ. Of decisive importance for the practical calculation is the approach of approximating the surface of the cavity using a finite number N_t of small, non-overlapping surface pieces, the so-called tesserae. In particular, they are selected in such a way that $\sigma(\vec{s})$ is an approximate constant for each tessera. Thus the surface charge is a discrete property, and a charge q_k can be introduced, which is comprised of the product of the local value of $\sigma(\vec{s})$ on a tessera, multiplied by the area of the tessera A_k. Thus the integral in equation (47) can be written as a finite sum:

$$V_R(\vec{r}) \approx \sum_{k=1}^{N_t} \frac{\sigma(\vec{s}_k) \cdot A_k}{|\vec{r} - \vec{s}_k|} = \sum_{k=1}^{N_t} \frac{q_k}{|\vec{r} - \vec{s}_k|} \tag{48}$$

The local value of the potential, which defines the charge q_k, itself depends on all charges on the surface. Charges and reaction field potential must consequently be determined iteratively. Naturally the values also depend on the selection of the discretization of the cavity surface. This will be explained in somewhat more detail in the section after next. At this point, we will merely insert a brief reflection. Through discretization according to (48), the solution of the Poisson equation is transformed into a matrix equation, which is often referred to as PCM equation. Such an approach shows a close relationship to the boundary element method (BEM), which is also referred to as the method of moments (MoM) in electro-technology. Here, a partial differential equation, which has a domain in $\mathbb{R}^3$, is solved numerically through integration of an integral equation with domain $\mathbb{R}^2$ as boundary surface. In BEM it is advantageous that the number of nodes is substantially lower than in the finite element method (FEM) or in the finite differences method (FDM). However, one does get a fully populated, nonsymmetric system of linear equations. A good introduction is provided by the monograph[28].

From (46) and additional boundary conditions (jump conditions), an integral equation can be derived that represents the PCM characteristic equation for $\sigma\left(\vec{s}\right)$[29]. Historically it was derived from elementary electrostatic considerations[27]. The polarization function of a medium i or of a spatially delineated region i with constant relative permittivity ε_{ri} is

$$\vec{P}_i\left(\vec{r}\right) = \frac{\varepsilon_{ri}-1}{4\pi}\vec{E} = -\frac{\varepsilon_{ri}-1}{4\pi}\nabla \mathrm{V}\left(\vec{r}\right) \tag{49}$$

At the boundary between two regions i and j, there is an ASC distribution of

$$\sigma_{ij} = -\left(\vec{P}_j\left(\vec{r}\right) - \vec{P}_i\left(\vec{r}\right)\right)\cdot \vec{n}_{ij} \tag{50}$$

where $\vec{n}_{ij}$ designates the unit vector on the boundary, which points from medium i to medium j. For systems described by CSMs (see also Figure 6), two regions are defined: inside the cavity with $\varepsilon_{r1} = 1$ and $\vec{P}_1 = \vec{0}$, outside the cavity with $\varepsilon_{r2} = \varepsilon_r$ and $\vec{P}_2 = \vec{P}$. Thus (50) can be re-written as a defining equation for PCM

$$\begin{aligned}\sigma\left(\vec{s}\right) &= \frac{\varepsilon_r - 1}{4\pi}\frac{\partial}{\partial \vec{n}}\left(V_M\left(\rho_M\left(\vec{r}\right)\right) + V_R\left(\sigma\left(\vec{s}\right)\right)\right)_{\mathrm{out}} \\ &= \frac{\varepsilon_r - 1}{4\pi\varepsilon_r}\frac{\partial}{\partial \vec{n}}\left(V_M\left(\rho_M\left(\vec{r}\right)\right) + V_R\left(\sigma\left(\vec{s}\right)\right)\right)_{\mathrm{in}}\end{aligned} \tag{51}$$

where $\vec{n}$ designates the unit vector that is perpendicular to the cavity surface and points outward. After discretization, as a defining equation for the charges q_k, one gets

$$q_k = \frac{\varepsilon_r - 1}{4\pi\varepsilon_r}\cdot A_k \cdot \frac{\partial}{\partial \vec{n}_k}\left(V_M + V_R\right)_{\mathrm{in}} \tag{52}$$

which must be solved iteratively and provides the self-polarized charges during convergence. Here, $\vec{n}_k$ designates a unit vector that is perpendicular to the surface of the cavity at point $\vec{s}_k$ and points outward (see Figure 7).

The problem can also be formulated as a system of linear equations, which provides advantages in implementation when linking to quantum mechanical methods.

$$\underline{\underline{T}}\vec{q} = \vec{V} \tag{53}$$

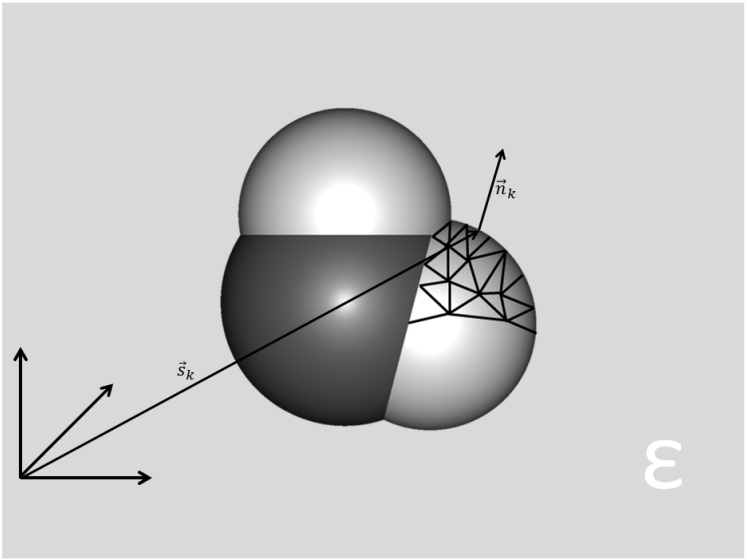


Figure 7: Sketch of the PCM for water with selected discrete quantities.

In (53) $\underline{\underline{T}}$ designates a quadratic nonsymmetrical matrix that is not singular and whose dimension corresponds to the number of tesserae. Here, the matrix elements are

$$\begin{aligned} T_{kk} &= \frac{4\pi\varepsilon_r}{\varepsilon_r - 1} \cdot (1 - B_k) \\ T_{kl} &= A_k \cdot \frac{\partial}{\partial \vec{n}_k} \frac{1}{|\vec{s}_k - \vec{s}_l|} = A_k \cdot \frac{(\vec{s}_k - \vec{s}_l)}{|\vec{s}_k - \vec{s}_l|^3} \cdot \vec{n}_k \end{aligned} \tag{54}$$

The factor B_k takes into consideration the self-interaction of the charges q_k and a geometric correction that is based on a specific construction of the cavity surface due to interlocking spheres (see also[27]).

$$B_k = \frac{\varepsilon_r - 1}{2\varepsilon_r} \cdot \left(1 - \sqrt{\frac{A_k}{4\pi R_k^2}}\right) \tag{55}$$

The radius of the respective interlocking sphere is R_k. Since the matrix $\underline{\underline{T}}$ depends only on the construction of the cavity surface and the permittivity, it can be inverted directly and stored. The vector $\vec{q}$ contains the charges, and the vector $\vec{V}$ contains the effective components of the electrostatic potential of the solute molecule at the positions $\vec{s}_k$.

$$V_k = -A_k \cdot \nabla V_M\left(\vec{s}_k\right) \cdot \vec{n}_k \tag{56}$$

A good summary, which also discusses some technical aspects, can be found in the review[30]; an original publication that deals with the different options for calculation of the apparent charges is reference[31] .

2.3.2 The quantum mechanical description

Since operator $\hat{V}$ in (39) for description of the solute-solvent interaction depends on the charge density of the solute itself, and this in turn depends on the reaction potential, described by this operator, which in turn acts back upon the charge density, a linked solution of the electrostatic and the quantum mechanical problem is necessary. Analogous to the quantum mechanical handling of the Onsager model, here too, an SCRF equation must be solved. The operator $\hat{V}_R$ is a function of the charge density ρ of the solute molecule M

$$\hat{V}_R(\rho) = \sum_{k=1}^{N_t} q(\vec{s}_k, \rho)\, \hat{V}(\vec{s}_k) \tag{57}$$

and the Schrödinger equation to be solved is defined by an effective Hamiltonian:

$$\hat{H}_{\text{eff}}\Psi = \left[\hat{H}_0 + \hat{V}_R\right]\Psi = \text{E}\Psi \tag{58}$$

The solute molecule in a vacuum is described by the usual nonrelativistic Hamilton operator. In the solvent, this is replaced by an effective Hamiltonian, which contains the self-consistent reaction field. The corresponding wave function Ψ depends on the mutual polarization of the solute and the solvent. Equation (58) can be derived from an energy functional. This functional corresponds thermodynamically to a free energy of the solute-solvent system. The reference state here is the solute molecule M at rest in the absence of an interaction of the nuclei and electrons and the undisturbed solvent at a given standard state for temperature and pressure.

$$\text{G} = \left\langle \Psi \,|\, \hat{H}_0 + \frac{1}{2}\hat{V}_R \,|\, \Psi \right\rangle \tag{59}$$

The factor $1/2$ results from the assumption of a linear dielectric and takes into consideration the work to be done in order to polarize the dielectric. The stationarity condition for (59) and the side condition of normalized wave functions $\langle \Psi \mid \Psi \rangle = 1$, lead to the Schrödinger equation (58) for the effective Hamilton operator, from which the electronic structure of the molecule can be determined. For an N-electron system, in which the spin orbitals are expanded in a basis of atomic orbitals (AO) $\{\chi_\alpha, \chi_\beta, \ldots\}$ (linear combination of atomic orbitals, MO-LCAO approach), the functional (59) can be written in the Hartree-Fock (HF) approximation as

$$G_{\text{HF}} = \sum_{\alpha\beta} P^{\text{HF}}_{\alpha\beta}\,(h_{\alpha\beta} + j_{\alpha\beta}) + \frac{1}{2}\sum_{\alpha\beta\gamma\delta} P^{\text{HF}}_{\alpha\beta} P^{\text{HF}}_{\gamma\delta}\,[\langle\alpha\gamma\|\beta\delta\rangle + X_{\alpha\beta,\gamma\delta}] + \tilde{V}_{\text{NN}} \tag{60}$$

Here, $h_{\alpha\beta}$ designates the matrix elements of the one-electron operator in the AO basis, $\langle\alpha\gamma\|\beta\delta\rangle$ the antisymmetrized two-electron integrals, and $P^{\text{HF}}_{\alpha\beta}$ the Hartree-Fock density matrix in the AO basis. The solute-solvent interaction is represented by the matrix elements $j_{\alpha\beta}$ and $X_{\alpha\beta,\gamma\delta}$. V_M in equation (41) can be represented as the sum of the solvent polarization by the charge distribution of the nuclei in solute M and by the electronic charge distribution of M: $V_M = V_M^n + V_M^e$. This partition also requires an analogous partition of the apparent charges. The matrix element $j_{\alpha\beta}$ is now defined as follows:

$$j_{\alpha\beta} \overset{\text{def}}{=} \vec{v}_{\alpha\beta} \cdot \vec{q}^{\,n} \tag{61}$$

where the components of the vector $\vec{v}_{\alpha\beta,}$ a N_t tuple, are the integrals of the electrostatic potential operator in the AO basis, each calculated at the positions $\vec{s}_k$ of the ASC charges on the corresponding tesserae, i.e. the matrix element

$$\left\langle \chi_\alpha \left| \frac{-1}{|\vec{r} - \vec{s}_k|} \right| \chi_\beta \right\rangle \tag{62}$$

$\vec{q}^{\,n}$ is the N_t tuple of the ASC charges based on the nuclear charge distribution. They are calculated by solving the system of linear equations (53), where for the vector on the right-hand side the potential generated by the nuclear charges

$$\sum_{\zeta=1}^{N_{\text{nuclei}}} \frac{Z_\zeta}{\left|\vec{R}_\zeta - \vec{s}_k\right|} \tag{63}$$

is used. Here, Z_ζ is an atomic number, $\vec{R}_\zeta$ the fixed position of a nucleus, and the index ζ runs via all N_{nuclei} nuclei of the solute molecule M. The potential (63) is not varied during the SCF iterations, and therefore the charges $\vec{q}^{\,n}$ can be determined at the start independently of the SCF calculation. The matrix element $X_{\alpha\beta,\gamma\delta}$ has the following form:

$$X_{\alpha\beta,\gamma\delta} \stackrel{\text{def}}{=} \vec{v}_{\alpha\beta} \cdot \vec{q}^{\,e}_{\gamma\delta} \tag{64}$$

where the $\vec{q}^{\,e}_{\gamma\delta}$ is determined via (53), where the N_t-component vector on the right side contains the electrostatic potentials (62) with $\alpha = \gamma$ and $\beta = \delta$. This contribution changes in each SCF cycle and must consequently be recalculated during the SCF iterations. Finally, there is a contribution to the nucleus-nucleus interaction, where, in addition to the nuclear repulsion V_{NN} (a constant factor in the Born-Oppenheimer approximation), it is necessary to take into consideration another contribution through the interaction of the nuclei and the polarization charges that are created by the nuclei themselves:

$$\tilde{V}_{\text{NN}} = \sum_{\zeta<\text{Z}}^{N_{\text{nuclei}}} \frac{Z_\zeta Z_\text{Z}}{\left|\vec{R}_\zeta - \vec{R}_\text{Z}\right|} + \frac{1}{2}\vec{v}^{\,n} \cdot \vec{q}^{\,n} \tag{65}$$

where the components of the N_t-component vector $\vec{v}^{\,n}$ are the potentials according to (63) at the positions $\vec{s}_k$ of the ASC charges. The requirement for the stationarity of the functional (60) $\delta G_{\text{HF}} = 0$ in regard to the variation of the MO-coefficients in satisfying the condition of ortho-normality results in the PCM-HF equations

$$\sum_\beta \left[f_{\alpha\beta}^{\text{PCM}} - \varepsilon_\gamma S_{\alpha\beta}\right] c_{\beta\gamma} = 0 \tag{66}$$

with the overlap matrix $S_{\alpha\beta}$ and the PCM Fock matrix $f_{\alpha\beta}^{\text{PCM}}$ each in the AO basis, and the orbital energy ε_γ , as well as the expansion coefficient $c_{\beta\gamma}$ for the MO γ. Here, the matrix elements of the Fock operator are:

$$f_{\alpha\beta}^{\text{PCM}} = (h_{\alpha\beta} + j_{\alpha\beta}) + \sum_{\gamma\delta} P_{\gamma\delta}^{\text{HF}} \left[\langle\alpha\gamma\|\beta\delta\rangle + X_{\alpha\beta,\gamma\delta}\right] \tag{67}$$

Proceeding from the functional (59), taking into consideration the definitions and formalism presented in outlining the Hartree-Fock method in PCM approximation, an analogous method can be formulated without any problem for the density functional theory (DFT).

There are corresponding implementations in a wide variety of program systems. By far the greatest number of publications on PCM are based on the DFT variant. Here, there are detailed and comparative studies on the influence of density functionals, the ways and means of constructing tesserae on the cavity surface, and on basis sets. Of very great importance for the usability of quantum chemical methods is the availability of gradients. There are now corresponding extensions for calculation of properties such as molecular geometries, infrared spectra, or chemical shifts. Good access to the literature is provided by the review[6]. More complicated than for HF and DFT, of course, is formulation of the PCM method on the level of the multiconfigurational self-consistent field (MCSCF) approximation. A pioneering work in this area is[32]. For many program systems that permit MCSCF calculations, especially the widespread programs, there are now implementations of PCMs. There are studies of various implementations of the PCM, taking dynamic correlation on the Møller-Plesset (MP) level into consideration. Since the electron correlation changes the total electronic charge distribution relative to the Hartree-Fock reference, the solvent reaction potential also varies. The modified polarization of the solvent in turn acts upon the effects of the electron correlation. Naturally, a quantitative examination of the individual effects is of interest, in order to permit a decoupled consideration and thus to get insights into the physical mechanisms of solvation. The first publications in this area introduced essentially three different variants of an implementation[33,34]. A noniterative variant (PTE), in which only the Hartree-Fock density on the PCM level is converged, and which is used to calculate the MPn correlation energies. The so-called density only scheme (PTD), in which the MPn density in a vacuum is used to calculate the reaction field. And finally the iterative variant (PTED), in which the correlated density is used to take the reaction field into consideration in a self-consistent manner. Numerical studies show a coupling of the different effects and thus demonstrate that, in the end, a consistent description is possible only on the PTED level[34,35]. There are now also implementations of the PCM on the coupled cluster (CC) level, the most precise method currently available for calculation of the electronic structure for molecules in a vacuum, which can be described well through a single-reference wave function. PCM CC implementations are available for the single and double excitation level (CCSD)[36,37] and on the Brueckner doubles level[38]. A study on the CC level on the influence of the reference wave function and orbital relaxation and the way to calculate the correlation density can be found in[39]. The literature on the state of the formalism and implementation of gradients in PCM CC is summarized in the review[40] .

2.3.3 The construction and the discretization of the cavity surface

As already discussed, an essential step in solving the electrostatic problem in PCMs is the discretization of a surface. Some aspects of this step will be outlined briefly here. Because of space constraints, technical details will be avoided. Part of the corresponding literature is available via the two reviews[4,6]. Furthermore, we refer to an review specifically on this topic[41]. In CSMs, the molecular cavity is a central concept. It designates a subset of the space in a medium, the solvent, which contains the solute molecule M. The boundary of this subspace is referred to as the molecular surface. The manner in which this cavity is defined and the way in which the molecular surface is constructed depends on the model. The first models had a very simple geometric form, e.g. a sphere or ellipsoid, for the molecular surface (see above). The most widespread manner of construction is based on

the easy-to-picture approach of placing a sphere around each atom position. The outer surface of the overlapping spheres then defines the molecular cavity (see also Figure 8).

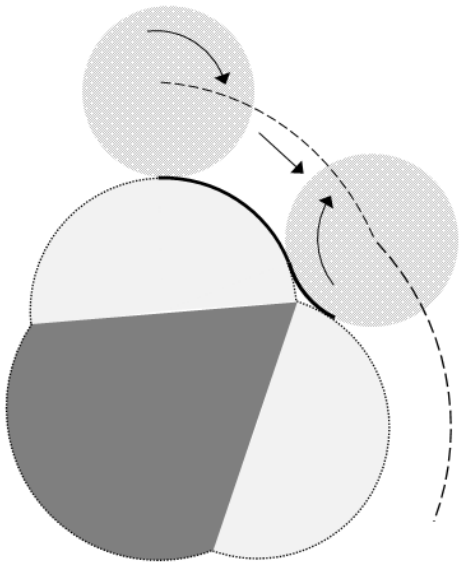


Figure 8: Sketch of van der Waals surface, solvent-accessible surface (dashed line), and solvent-excluded surface (bold line).

This manner of construction can be specified further. Here, one distinguishes essentially among three types of molecular surfaces:

1. The van der Waals surface (VWS) is the surface that results from the construction of spheres whose radius corresponds to the corresponding van der Waals radius. Corresponding radii are found in reviews; the "classic" being the publication by Bondi[42].

2. The solvent-accessible surface (SAS) is constructed from the VWS, by rolling a sphere (the solvent sample) with a prescribed radius over the VWS and selecting as molecular surface the number of points created by the middle point of the moving spheres. The radius of the sample can be adapted to the characteristics of the solvent. It can be created simply by adding the radius of the solvent to each van der Waals radius.

3. The solvent-excluded surface (SES) is defined as the number of contact points that a solvent sample, defined by a sphere of a given radius, has with a VWS. Through this process, certain areas of very small volume and complex geometry (e. g. narrow gaps) are "wiped away". In comparison to the VWS, the surface is smoothed. The intersection of VWS and SES is now defined as those points that correspond to the convex part of the SES. The remaining ones, whose points are not elements of the

VWS, correspond to concave surfaces. The space encompassed by the SES, but not by the VWS, is referred to as the solvent-excluded volume.

In contrast to VWS and SAS, which can be constructed from spheres, the construction of the SES is more complicated. At this point, because of the limited space, details will not be presented; however, we can refer to a classic work by Connolly, which, both conceptionally and practically, is of great significance far beyond the topic of CSMs[43]. If there is a molecular surface, this surface must be divided into small areas in a further step. This defines the tesserae, from which the set of the vectors $\vec{s}_k$ and their components can be calculated (see Figure 7). For this purpose, a number of algorithms have been developed and implemented over the course of the development of the PCM. For details and good access to the literature, refer to the review[41]. Here we will merely mention the name of the algorithm with which the development of the PCM was intimately linked, particularly in the early days: GEPOL[44] (although the method has already been widely used, publications in scientific journals did not come until much later).

2.3.4 The conductor-like screening approach

The complicated electrostatic problem of solving the Poisson equation for complex, but in the broadest sense realistic molecular cavities as discussed in the previous section can be drastically simplified, by assuming in the PCM, not a solvent initially characterized by a finite relative permittivity, but one that is infinitely large. Thus it is assumed that the solute molecule is located in a perfect electrical conductor. The surface charges are calculated for the boundary condition that the electrostatic potential on the surface Γ is zero. This gives this approach its name, conductor-like screening model, which was introduced in reference[45]. The matrix and vector elements of the linear equation system (53) for determination of the ASC now consist of simple properties already present in every elaborated quantum chemistry program.

$$\underline{\underline{\tilde{T}}}\vec{q}^{\,*} = \vec{\tilde{V}} \tag{68}$$

The matrix elements of $\underline{\underline{\tilde{T}}}$ are:

$$\begin{aligned} \tilde{T}_{kk} &= const. \cdot \sqrt{\frac{4\pi}{A_k}} \\ \tilde{T}_{kl} &= \frac{1}{|\vec{s}_k - \vec{s}_l|} \end{aligned} \tag{69}$$

and the elements of the vector $\vec{\tilde{V}}$ consisting of N_t components are the matrix elements (62). Since the charges calculated according to (68) constitute the limiting case $\varepsilon_r \longrightarrow \infty$ they must be corrected for the case of finite permittivity:

$$\vec{q} = f\left(\varepsilon_r\right) \cdot \vec{q}^{\,*} \tag{70}$$

$f\left(\varepsilon_r\right)$ is[45]:

$$f(\varepsilon_r) = \frac{\varepsilon_r - 1}{\varepsilon_r + k} \quad \text{with} \quad 0 \leq k \leq 1 \tag{71}$$

The proposal of the original work was $k = 0.5$[45]. This value is not without controversy. The review[6] provides good access to the corresponding literature. The COSMO approach can be implemented very efficiently in quantum chemical programs. In particular in calculating gradients, this form of PCM is advantageous. Details on this topic can be found in the publications[46,47] .

2.3.5 Documentation of important apparent surface charge methods

Proceeding from the integral representation of the reaction field potential (47), after introducing certain integral operators, the PCM variants can be written in the form of integral equations (regarding an introduction, refer to the corresponding section of the review[6]). This new formalism (Integral Equation Formalism, IEF) not only provided an elegant, uniform representation, but also made it possible to find a number of additional PCM variants. The most important of them will now be documented in the form of a brief summary. Viewed from the perspective of calculations, the methods differ basically in the defining equation for the ASCs. For a summary presentation, analogous to (53), the system of linear equations is formulated as[6]:

$$\vec{q} = -\underline{\underline{K}} \cdot \vec{f} \tag{72}$$

A number of definitions for matrix elements are provided in the following:

$$\begin{aligned} A_{kk} =& A_k; A_{kl} = 0 \\ S_{kk} = const. \cdot& \sqrt{\frac{4\pi}{A_k}}; S_{kl} = \frac{1}{|\vec{s}_k - \vec{s}_l|} \\ D_{kk} = D^*_{kk} =& const. \cdot \frac{1}{R_k}\sqrt{\frac{\pi}{A_k}} \\ D_{kl} = -\frac{(\vec{s}_k - \vec{s}_l)}{|\vec{s}_k - \vec{s}_l|^3} \cdot \vec{n}_l; & D^*_{kl} = -\frac{(\vec{s}_k - \vec{s}_l)}{|\vec{s}_k - \vec{s}_l|^3} \cdot \vec{n}_k \end{aligned} \tag{73}$$

The approach described in 2.3.1, the historically first CPM variant, is referred to as "DCPM" in the most recent literature, in order to distinguish the other variants. As already mentioned above, R_k designates the radius of the sphere to which the tessera k belongs. Depending on implementation, the constants can differ. For D_{kk},, D^*_{kk}, and D^*_{kl}, there are also definitions that deviate from (73). Please refer to the original literature; it may sometimes also be necessary to consult the corresponding technical documentation on the individual quantum chemical programs.

1. DPCM

$$\underline{\underline{K}} = \left(2\pi\frac{\varepsilon_r + 1}{\varepsilon_r - 1}\underline{\underline{A}}^{-1} - \underline{\underline{D}}^*\right)^{-1}; \vec{f} = \vec{E}_n \tag{74}$$

2. COSMO, also referred to as CPCM

$$\underline{\underline{K}} = \underline{\underline{S}}^{-1}; \vec{f} = \vec{V} \tag{75}$$

3. IEFPCM, also named IVPCM

$$\underline{\underline{K}} = \left[\left(2\pi\frac{\varepsilon_r+1}{\varepsilon_r-1}\underline{\underline{A}}^{-1} - \underline{\underline{D}}\right)\underline{\underline{S}}\right]^{-1}\left(2\pi\underline{\underline{A}}^{-1} - \underline{\underline{D}}\right); \vec{f} = \vec{V} \tag{76}$$

4. SS(V)PE (acronym for "surface and simulation of volume polarization for electrostatics")

$$\underline{\underline{K}} = \left(2\pi\frac{\varepsilon_r+1}{\varepsilon_r-1}\underline{\underline{A}}^{-1}\underline{\underline{S}} - \underline{\underline{A}}^{-1}\frac{\left(\underline{\underline{D}}\cdot\underline{\underline{A}}\cdot\underline{\underline{S}} + \underline{\underline{S}}\cdot\underline{\underline{A}}\cdot\underline{\underline{D}}^{*}\right)}{2}\right)^{-1}\left(2\pi\underline{\underline{A}}^{-1} - \underline{\underline{D}}\right); \vec{f} = \vec{V} \tag{77}$$

Here, $\vec{V}$ refers to the N_t tuple of matrix elements of the electrostatic potential multiplied with -1 (62); and $\vec{E}_n$ refers to the corresponding vector elements for the flux of the electrostatic field through the tessera k.

2.3.6 The problem of escaping (outlying) charges

Fundamental to the solution of the electrostatic problem in PCM is the assumption that the charge distribution ρ_M is contained completely within the cavity. Only under this prerequisite is the problem defined by the equations (44) to (46). For quantum mechanical calculations, however, this is fundamentally not the case, because the wave function of the solute always has tails, which extend across the boundary of the cavity into the space of the solvent. Formally, it is relatively easy to develop a corresponding modification of the electrostatic problem. However, formulating and executing corresponding numerical calculation methods does become a problem. If the potential of the solute-solvent interaction (41) is extended by one term, the effects of escaping charges are taken into consideration.

$$V\left(\vec{r}\right) = V_M\left(\vec{r}\right) + V_R\left(\vec{r}\right) + \frac{1}{\varepsilon_r}V_{M,ext}\left(\vec{r}\right) \tag{78}$$

The additional potential in (78) is defined based on (43):

$$V_{M,ext}\left(\vec{r}\right) = \int_{\mathbb{R}^3\setminus C} d\vec{\tilde{r}}\,\frac{\rho_M\left(\vec{\tilde{r}}\right)}{|\vec{r}-\vec{\tilde{r}}|} \tag{79}$$

The integration in (79) must be done throughout the entire volume, with the exception of the molecular cavity. This is no simple numerical problem. Furthermore it is not obvious how to determine the potentials V_M and $V_{M,ext}$ numerically in quantum chemical programs. In practice the problem of the outlying charges is therefore either simply ignored or corrected empirically. The latter will not be discussed in any further detail here. However, we will briefly outline how the outlying charge error can be quantified, and a basic estimate will be provided for the sensitivity of the PCM variants discussed here in regard to this deficit. The total charge of the solute-solvent entity is defined via the integral

$$Q = \int_{\mathbb{R}^3} d\vec{r}\rho_M\left(\vec{r}\right) \tag{80}$$

The escaped charge can then be estimated via the relationship

$$Q_{esc} = Q - \int_C d\vec{r}\rho_M\left(\vec{r}\right) = Q + \frac{1}{4\pi}\int_C d\vec{r}\Delta V_M\left(\vec{r}\right) = Q + \frac{1}{4\pi}\int_\Gamma d\vec{s}\frac{\partial}{\partial\vec{n}}V_M\left(\vec{s}\right) \tag{81}$$

The second identity follows due to (45) and the third through application of Gauss's theorem. If the ratio Q_{esc}/Q is on the order of a several per cent, then one can conclude that there was a defect in the calculation. Analytical examinations of the integral equations of the PCM variants led to fundamental conclusions as to the quality of the approximations regarding the outlying charges. For instance, in[48] it was shown that for the exact reaction field energy

$$E_R(\rho, \tilde{\rho}) = \int_{\mathbb{R}^3} d\vec{r}\tilde{\rho}(\vec{r}) V_R(\vec{r}) \tag{82}$$

and that calculated with the IEFPCM method $E_R^{\mathrm{IEFPCM}}(\rho, \tilde{\rho})$, the following inequality condition applies:

$$E_R(\rho, \tilde{\rho}) \leq E_R^{\mathrm{IEFPCM}}(\rho, \tilde{\rho}) \tag{83}$$

The IEFPCM method thus provides an upper limit for the exact reaction field energy. The electrostatic problem for the potential as per equation (78) was studied in a number of publications (for a summary, see e.g.[49]) and the solution is known under the name "surface and volume polarization for electrostatics" (SVPE). This method is complicated, because, in contrast to the previously discussed PCM methods, here it is necessary to discretize a volume polarization potential that is created by a discontinuous charge density. Proceeding from this approach, it was possible to derive a new ASC method: SS(V)PE[50]. Put roughly, one can say that the IEFPCM and the SS(V)PE method, but also COSMO, do better with the outlying charges than does DPCM. In the end, however, such an assessment depends on the individual case.

3 Some Illustrative Data on Calculations with Polarizable Continuum Models

Based on the very simplified description of solvation by CSM and PCM approaches, a cautious approach to such methods is unavoidable. In the end, great errors must be expected, and it is therefore indispensable before conducting a study with PCMs to do benchmarking for comparable systems, so that an assessment of the error bars can be done. CSM and PCM approaches contain empirical parameters. They are specific for each implementation, which makes the situation confusing. Statements that model X is better than model Y should be taken with a grain of salt, because, the deficit of the one model can be related to the fact that the parameters have simply not been adapted well to the particular molecule class being examined. For another class of molecules, however, this can certainly be different. In addition to the consideration or non-consideration of the contributions to solvation free energy, which are not of an electrostatic nature (see equation (3)), and their specific implementation, we should also mention in particular the atomic radii and the nature of the cavity construction. As a rule the parameters are selected in such a way that a particular training data set of solvation free energies is reproduced accurately at 298 K.

In the following, certain data from a recently published, comparative study on the precision of PCM methods in the calculation of the hydration energy of the amino acid side chain analogs[51] are reviewed and commented. The values documented here were calculated by doing a geometry optimization in a vacuum for each of the 14 molecules. Then, for each, a single-point calculation was done in a vacuum (if the method or the basis set

deviated from that of the geometry optimization) and with PCM methods with the same quantum chemical method and basis set. The difference of these two energy values was assumed as the hydration free energy. In the study, samples examinations were done to determine how greatly the geometries in a vacuum deviate from those with PCM methods and what influence this has on the hydration free energy. These effects were very small and can therefore be disregarded. The experimental data at 298 K for the 14 compounds cover an interval of -10.3 to +2.3 kcal/mol. If one examines the calculations with UAKS radii, it can be seen that the trend of the experimental data is reproduced very well. The maximum deviation is under 1 kcal/mol. This applies for all the combinations of quantum chemical methods and PCM presented here. Taking correlation effects into consideration leads to a somewhat better description. Considering the approximations inherent in the models, there is no reason to discuss any further detailed assessments of the systems discussed here.

Method	Atomic radius	Mean abs. dev. kcal/mol	Stand. dev. kcal/mol
HF/6-31G**/IEFPCM	UAKS	0.35	0.27
	Bondi	1.09	0.72
	UFF	5.13	2.14
MP2/aug-cc-pVTZ/IEFPCM	UAKS	0.25	0.22
	Bondi	1.12	0.73
	UFF	5.67	2.43
DFT-B3LYP/6-31G**/IEFPCM	UAKS	0.32	0.27
	Bondi	1.23	0.79
	UFF	5.87	2.51
DFT-B3LYP/6-31G**/CPCM	UAKS	0.29	0.25
DFT-B3LYP/6-31G**/DPCM	UAKS	0.27	0.25

Table 1: Data on calculations with PCM for the hydration free energy of the 14 compounds as per[51]: methane, propane, iso-butane, n-butane, methylethylsulfide, toluene, para-cresol, 4-methylimidazole, 3-methylindole, methanethiol, methanol, ethanol, acetamide, propionamide

The dependence of the quality of the results upon the atomic radii employed, which are used for construction of the molecular cavity, is very striking and typical. Independent of the PCM used (note that the non-electrostatic contributions to solvation free energy are identical in the methods used here) or quantum chemical level, they have the decisive influence upon the quality of the results.

Finally, typical data on the calculation of free energies of chemical reactions will be discussed, in order to illustrate how PCMs can be used for some of the most important questions to computational chemistry. Prototypical for a significant class of reactions in organic chemistry is the esterification of carboxylic salts with alkyl halides. From a publication on calculations on the model reaction

$$CH_3COO^- + CH_3CH_2X \xrightarrow{S_N2} CH_3COOCH_2CH_3 + X^- \tag{84}$$

in which the geometry optimizations on the MP2 level and the single-point calculations for energy on the MP4 level were calculated with the DPCM model (refer to the original

work for basis sets and other technical details)[52]; for the individual species when using chlorides, bromides, and iodides in the two solvents dimethylsulfoxide and water, the free energies are documented in the following table.

Species	$\Delta G_{\mathrm{solv}}^{\mathrm{PCM}}$ (DMSO)	$\Delta G_{\mathrm{solv}}^{\mathrm{exp}}$ (DMSO)	$\Delta G_{\mathrm{solv}}^{\mathrm{PCM}}$ (H_2O)	$\Delta G_{\mathrm{solv}}^{\mathrm{exp}}$ (H_2O)
CH_3COO^-	-64.11	-62.7	-75.87	-77.3
Cl^-	-64.21	-65.0	-73.58	-74.6
CH_3CH_2Cl	-1.91		-3.40	
$TS-Cl$	-50.63		-61.87	
Br^-	-61.10	-62.1	-67.79	-68.6
CH_3CH_2Br	-1.99		-3.42	
$TS-Br$	-49.30		-59.46	
l^-	-55.75		-60.86	
CH_3CH_2l	-1.87		-3.27	
$TS-l$	-47.15	-57.4	-56.76	-59.9
$CH_3COOCH_2CH_3$	-4.05		-8.16	

Table 2: Calculated and experimental free solvation energies in kcal/mol[52]. $TS-X$ designates the transition state.

The data make it clear that the selected combination of quantum mechanical method and PCM is able to provide quantitative predictions, in particular when one takes into consideration that the comparisons with the experimental data apply for the ionic systems, which, particularly for reactions between ionic system and neutral species, are as a rule, the more difficult to describe components. The errors documented here are typical for a good description of the solvation free energy. As a rule, no better agreements can be expected. It should also be noted here that comparative studies have demonstrated well that for S_N2 reactions of the type (84), DFT methods are not suitable, particularly for the quantitative calculation of reaction barriers, and that one has to use ab initio methods with good determination of the correlation energy. Table 3 documents the calculated barriers and those derived from experimental studies (for details refer to[52]).

The data show that calculations on this level can reflect well the reactivities for reactions of the type (84) in various solvents. This applies analogously for the thermochemical data, which, however, will not be discussed further here. Thus it is possible to do a purely computational analysis of technical questions involving reactions and, in particular, those regarding the selection of a solvent. A variety of reactions in solution can be described in similar quality by PCMs on the DFT level. Thus reactions with molecules containing on the order of 1000 atoms in solution can be studied, which also means that a significant subset of industrially relevant reactions, in particular including homogeneously catalyzed reactions, are accessible to purely computational description.

$R-X$	$\Delta G^{\#}$ (calc)	$\Delta G^{\#}$ (exp)	deviation
in DMSO			
$R-Cl$	24.9	22.3	2.6
$R-Br$	20.0	20.0	0.0
$R-I$	18.5	16.6	1.9
in H_2O			
$R-Cl$	26.9	26.1	0.8
$R-Br$	23.1	25.2	-2.1
$R-I$	22.1	24.7	-2.6

Table 3: Calculated and experimental reaction barriers for the reaction (84) as per[52] in kcal/mol.

4 The Conductor-Like Screening Model of Real Solvents (COSMO-RS)

To conclude these lecture notes a brief overview of an approach shall be given which uses properties from PCMs, and in particular a quantity derived from the ASCs, to gain access by a statistical thermodynamic approach to calculation of the central property of mixed-phase thermodynamics, which is the chemical potential. This method is named "conductor-like screening model of real solvents" (COSMO-RS)[2]. A source for the literature up to 2004 is the monograph[53]. A subset of the literature from 2005 onward is summarized in the review article[54]. A brief summary of a subset of important methodological aspects is available[55].

4.1 The COSMO-RS model for interaction energy in solutions

The starting point of the approach is the assumption that the solute molecules are in a structureless medium that has the electrostatic properties of a perfect conductor. Technically speaking, therefore, molecules can be described using COSMO for the limiting case $\varepsilon_r \to \infty$. The electric field of the molecules is in every case perfectly shielded by a surface polarization charge density. In the event of the molecules being localized at large distances from one another, no mutual interactions exist between them. The molecules are characterized in every case by surface charge densities. This situation is the reference state in COSMO-RS. If molecules come sufficiently close to one another, mutual interaction occurs between them. An important contribution to the interaction energy is based on electrostatics. This interaction is described by the assumption in the model that small, localized, charged areas exist, which come into contact. To arrive at a quantitative description, it is postulated that these local pairwise interactions of the polarization charge densities σ and $\tilde{\sigma}$ result in a total charge leading to a so-called misfit energy contribution that depends on the square of this total charge.

$$\Delta E_{\text{misfit}} = \alpha \cdot (\sigma + \tilde{\sigma})^2 \tag{85}$$

From empirical observation (see[56]), after application of the COSMO-RS model (of which further details are developed below) based only on (85) it is found that an additional term,

which explicitly describes hydrogen bonding, is necessary for the description if the method developed is to have quantitative predictive power. This term, which is chosen ad hoc, is of the form:

$$\Delta E_{\mathrm{HB}} = c_{\mathrm{HB}} \cdot \min\left(0, \min\left(\sigma, \tilde{\sigma}\right) + \sigma_{\mathrm{HB}}\right) \cdot \max\left(0, \max\left(\sigma, \tilde{\sigma}\right) - \sigma_{\mathrm{HB}}\right) \qquad (86)$$

where $\min$ and $\max$ are the minimum and maximum functions from order theory that supply the smallest and the largest element of an ordered set. The phenomenon of hydrogen bonding is thereby reduced to being described by surface charges, it being assumed here that a certain threshold σ_{HB} must be attained for a charged area to be able to form a hydrogen bond. The sum of (85) and (86) now defines the interaction energy $\Delta E_{\mathrm{int}}\left(\sigma, \tilde{\sigma}\right) \stackrel{\mathrm{def}}{=} \varepsilon\left(\sigma, \tilde{\sigma}\right) \stackrel{\mathrm{def}}{=} \Delta E_{\mathrm{misfit}} + \Delta E_{\mathrm{HB}}$ in COSMO-RS.

4.2 The COSMO-RS integral equation

The central equation of COSMO-RS is obtained from a model concept of interactions in liquids that represents an extreme simplification of the complex statistical thermodynamics. It is assumed that all interactions in a liquid, i.e., in an ensemble S of n different molecules $M_i \,|i\epsilon \left\{1, 2, \cdots, n\right\}$, can be described by taking into account only the pairwise interactions of surface areas. This neglects, in particular, all information of a geometric nature. An in a certain sense similar approach, which is not obvious from an atomistic viewpoint, has long been very successfully used for predictive description of mixed-phase thermodynamic properties in chemical engineering. It is based essentially on the fundamental work of Guggenheim (see, for example, the monograph[57]). These so called group contribution methods are based to a very large extent on experimental data, and have been routinely used for many years. Process development in the chemical industry is now inconceivable without recourse to such methods. An overview of these methods is provided in the monographs[58,59]. The relationship between modern methods of this class and COSMO-RS has been worked out in[60]. From a statistical thermodynamic theory for pairwise interacting surface charges, the following integral equation (which in the present model is exact) can be derived to determine the chemical potential as a function of the surface screening charge[2,61]

$$\mu_S\left(\sigma\right) = -\frac{RT}{a_{\mathrm{eff}}} \ln\left[\int_a^b d\tilde{\sigma}\, P_S\left(\tilde{\sigma}\right) \exp\left\{a_{\mathrm{eff}} \cdot \frac{-\varepsilon\left(\sigma, \tilde{\sigma}\right) + \mu_S\left(\tilde{\sigma}\right)}{RT}\right\}\right] \qquad (87)$$

The property $P_S\left(\sigma\right)$ is what is known as the σ-profile of the ensemble S. It is composed of the σ-profiles of the molecules in the following way:

$$P_S\left(\sigma\right) = \sum_{i=1}^{n} x_i \cdot p^{M_i}\left(\sigma\right) \qquad (88)$$

Here x_i is the mole fraction of molecule i and $p^{M_i}\left(\sigma\right)$ the so-called σ-profile of molecule i. The integral equation (87) belongs to the class of Hammerstein integral equations[62]. It can be solved either numerically after discretization[2] or semi-analytically[62]. In the course of essentially empirical studies, it was found that it is vitally important not to directly use the ASCs and the associated surfaces from a COSMO calculation. Instead, a new surface charge is calculated from the ASCs, obtained by averaging over larger areas (the details

are available in[56]). For purely numerical solution of (87) this equation is self-consistently solved for each of the charges, to determine the chemical potential as a function of the surface charge. From these surface charges obtained by averaging, the σ-profiles are also determined. The σ-profiles are very interesting quantities because they provide a clear, detailed, and immediate fingerprint of the electrostatic properties. The fingerprint contains, in the condensed form of a histogram, information on the screening charge density at the surface of the cavity of the molecule. From this it is possible to infer, for example, what proportion of the surface of the cavity of the molecule lies within a given interval of screening charges. Figure 9 shows the σ-profiles of a few molecules.

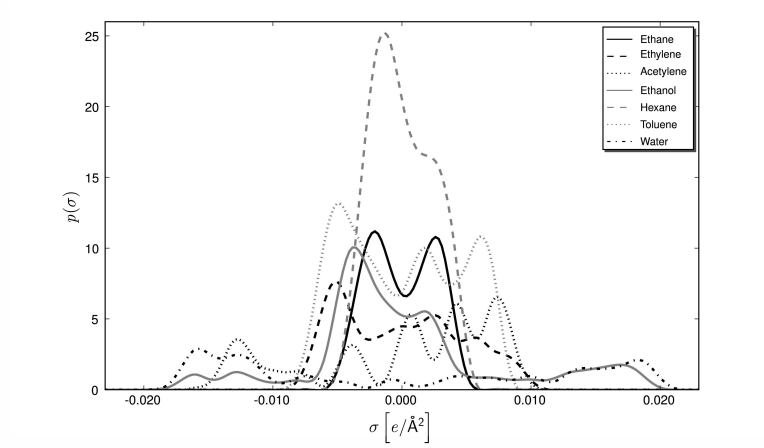


Figure 9: Selected σ-profiles

We consider the case of water. As can be seen, the σ-profile is relatively broad and symmetric. It shows two maxima: at $-0.016e/\text{Å}^2$ and at $+0.018e/\text{Å}^2$. The latter, at positive screening charge density, corresponds to the lone pairs of the partially negatively charged oxygen atom, while the former, at negative screening charge density, corresponds to the two partially positively charged hydrogen atoms. Of note here is that in σ-profiles negative partial charges always correspond to positive screening charge densities and vice versa. The shape of water's σ-profile reflects the ability to act as a hydrogen bond donor as well as an acceptor. We shall not go into a more detailed explanation of their form here; detailed interpretations are available in, for example,[53]. The integration limits of the integral on the right-hand side of equation (87) are obtained empirically by consideration of shapes for $p^{M_i}(\sigma)$. $\mu_S(\sigma)$ is termed the σ-potential. It is a function of the composition of the liquid mixture and of the temperature. This property is also a fingerprint of the behavior in liquid mixtures. It basically describes the affinity of a solvent, which may be composed of many different molecules in a complex way, for a dissolved substance. For discussion of specific σ-potentials see[53].

4.3 The chemical potential

The chemical potential $\mu_S^{M_i}$ of a molecule M_i in an ensemble S of any arbitrary complexity (at least in principle) is obtained as a simple integral over σ-profile multiplied by σ-potential as well as an additional term that essentially accounts for entropic effects.

$$\mu_S^{M_i} = \int_a^b d\sigma \, p^{M_i}(\sigma) \cdot \mu_S(\sigma) + \mu_{CS}^{M_i} \tag{89}$$

The integral on the right-hand side of equation (89) denotes the so-called residual part of the chemical potential. The second term denotes the combinatorial part of the chemical potential. This breakup of the chemical potential follows chemical engineering terminology. For $\mu_{CS}^{M_i}$ a number of empirical formulas exist which incorporate the surface area and volume, inter al., of the molecules. A more detailed discussion will not be offered here; the reader is referred instead to[58,53]. The proportion is generally small compared with the residual part, but cannot be neglected for a quantitative description. The central property of mixed-phase thermodynamics is thus available in principle for liquid phases of any composition. Liquid-liquid equilibria (LLE) can thus be described almost without any restriction. For this purpose only computations for the individual molecules M_i at the COSMO level with $\varepsilon_r \to \infty$ are required. Through the equations for the COSMO-RS method shown here, the ASCs allow access to the chemical potential for a mixture of given composition at a given temperature. We will once again mention here the important property of COSMO-RS that the calculations for determining the ASCs with COSMO, which are by far the most computer-time-intensive step of this method, need be performed only once. (The solution of equation (87) can be very efficiently implemented; the computer time required for this is basically negligible.) The ASCs are typically stored in a database so as to have them available for later COSMO-RS calculations. With this method, therefore, it is possible to determine, very efficiently and without recourse to experiment, the substance data for design of chemical processes based on mixed-phase thermodynamic properties. This is the reason for the great popularity of this method in the areas of chemical engineering and industrial chemistry. LLEs are of course also overwhelmingly important in the life sciences. The preferred fields of application here are solvent screening and formulation research in drug development. The above-mentioned parameters in the equations for COSMO-RS $(\alpha, c_{\mathrm{HB}}, \sigma_{\mathrm{HB}}, a_{\mathrm{eff}})$, as also the atomic radii for the underlying COSMO calculations and the parameters for the combinatorial part of the chemical potential as well as certain other parameters not explicitly mentioned here, are obtained by fitting of experimental data sets. This is therefore a semi-empirical method. Differences of chemical potentials can be reproduced with an accuracy of 0.5 kcal/mol with a global parameterization that claims to apply for all neutral compounds in the liquid phase. This accuracy is adequate in a large number of cases. The level of quantum chemical description—the dependence of the functional used in calculations with DFT and COSMO, for example, or the basis set used—play a subordinate role in the quality of a COSMO-RS parameterization[63]. This has its advantages, but also certain disadvantages. It means, on the one hand, that computing expenditure on the quantum chemical side can be kept low. There is little hope, however, that accuracy will be increased significantly by an improved description at the quantum-chemical level. This would be highly desirable because the best empirical group contribution methods routinely used in process simulation show significantly smaller error bars. Their use is of course restricted to those systems for which reliable experimental data are available for their parameterization. One approach toward improving the quality of

prediction with COSMO-RS almost to the level of the purely empirical group contribution methods consists in a parameterization of COSMO-RS for only those molecule classes relevant for a given problem[64]. The resulting loss of generality of the parameterization often presents no problem in practical applications.

Reference[54], for example, gives an idea of the varied and successful applications of COSMO-RS that have been documented in the literature. The model naturally has many limitations: dynamic properties like viscosity or diffusion coefficient, and properties in the region of or beyond the critical point cannot be described. Nor is the solvation of ions like F^- or Al^{3+} well described. Moreover, there are certain classes of molecules, such as tertiary amines, for which COSMO-RS shows particularly large errors.

5 Concluding Remarks

The lecture notes provide a brief introduction to the theory of CSMs and its application in the area of PCMs. This is a class of methods that are used more or less as the standard for description of solvation effects when using quantum chemical methods. These methods are easily available and, in comparison to a calculation in a vacuum, require only a little additional calculation effort. If their inherent approximations and thus the limits of the model are taken into consideration along with careful prior examination of their suitability in model systems, which are adequately similar to those of the problem being worked on, it is possible to make quantitatively and qualitatively good predictions. A semi-empirical, statistical-thermodynamic extension of the COSMO approach, COSMO-RS, permits calculation of the chemical potential in complex mixtures at variable temperatures. Thus, in principle, this provides access to the complete range of mixed phase thermodynamics in liquid phase. CSMs are based on a drastically simplified description of solvation, in which it is reduced to the modeled handling of the solvent in a homogeneous, isotropic, linear dielectric medium. It is assumed that the solvent is located in a cavity in this medium. The degree to which the dielectric is polarized and its associated ability to create an electric field (the reaction field) that in turn acts back upon the solvent, is determined by a macroscopic property, the relative permittivity (static dielectric constant). The charge distribution of the solute is now optimized self-consistently in the presence of the reaction field (self-consistent reaction field calculation, SCRF). The calculation is to be done iteratively, because the reaction field itself again depends upon the modified solute wave function. The description of the influence of the solvent by a homogeneous, isotropic, linear dielectric is justified only for solvent molecules that are far away from the solute. Particularly large errors are to be expected for the description of the solute-solvent interaction in the first solvation shell. Here discrete, directed interactions dominate, which depend upon local details of the charge distribution on the atomistic scale. When the structural characteristics of the liquid are taken into consideration, the linear response approximation no longer applies. The polarization here is of a nonlinear nature. The spatial division into a space containing the solute molecule and a sharply separated space for the dielectric medium is arbitrary and basically impermissible. Free energies resulting as solutions to electrostatic model problems based upon such assumptions are therefore always somewhat arbitrary. In the end, the undeniable success of such models in describing certain solvation phenomena and in making predictions is based upon empirical adjustment. This must always be taken into consideration when applying methods of this class and in assessing published numerical values based upon such calculations.

References

1. R. Di Remigio, R. Bast, L. Frediani, T. Saue. J. Phys. Chem. A, 2014, DOI:10.1021/jp507279y.
2. A. Klamt, J. Phys. Chem., **99**, (1995), 2224-2235.
3. P. Deglmann, A. Schäfer, C. Lennartz, Int. J. Quantum Chem., **115**, (2015), 107-136.
4. J. Tomasi, M. Persico, Chem. Rev., **94**, (1994), 2027-2094.
5. C. J. Cramer, D. Truhlar, Chem. Rev., **99**, (1999), 2161-2200.
6. J. Tomasi, B. Mennucci, R. Cammi, Chem. Rev., **105**, (2005), 2999-3093.
7. B. Mennucci, R. Cammi. *Continuum Solvation Models in Chemical Physics: From Theory to Applications*, John Wiley & Sons Ltd, Hoboken, 2007.
8. N. H. List, H. J. A. Jensen, J. Kongsted, E. D. Hedegaard, Adv. Quantum Chem, **66**, (2013), 195-238.
9. M. Born, Z. Phys. **1**, (1920), 45-48.
10. W. Pauli, *Pauli Lectures on Physics*, vol. 1, The MIT Press, Cambridge, 1973.
11. G. Ludwig, *Einführung in die Grundlagen der Theoretischen Physik*, Band 2, Verlagsgruppe Bertelsmann, Düsseldorf, 1974.
12. Jackson, J. D. *Classical Electrodynamics*, 2nd edn, John Wiley & Sons, New York, 1975.
13. W. Nolting, *Grundkurs Theoretische Physik*, Band 3, Springer-Verlag, Berlin, 2011.
14. F. Figueirido, G. S. Del Buono, R. M. Levy. Biophys. Chem., **51**, (1994), 235-241.
15. M. Luckas, J. Krissmann, *Thermodynamik der Elektrolytlösungen*, Springer-Verlag, Berlin, 2001.
16. L. Onsager, J. Am. Chem. Soc., **58**, (1936), 1486-1493.
17. P. Debye, Physik. Z., **13**, (1912), 97-100.
18. J. Tomasi, Theor. Chem. Acc., **103**, (2000), 196-199.
19. C. J. F. Böttcher, *Theory of Electric Polarization*, vol. 1: Dielectrics in static fields, 2nd edn, Elsevier, Amsterdam, 1973.
20. M. W. Schmidt, T. L. Windus, M. S. Gordon. J. Am. Chem. Soc., **117**, (1995), 7480-7486.
21. O. Tapia, O. Goschinski. Mol. Phys., **29**, (1975), 1653-1661.
22. R. J. Abraham, E. Bretschneider. *Medium Effects on Rotational and Conformational Equilibria* in: W. Orville-Thomas (ed.), Internal Rotational in Molecules, John Wiley & Sons, London, 1974.
23. M. W. Wong, M. J. Frisch, K. B. Wiberg. J. Am. Chem. Soc., **113**, (1991), 4776-4782.
24. Y. Luo, H. Agren, K. V. Mikkelson. Chem. Phys. Lett., **275**, (1997), 145-150.
25. K. V. Mikkelsen, H. Agren, H. J. A. Jensen, T. Helgaker, J. Chem. Phys., **89**, (1988), 3086-3095.
26. A. Milischuk, D. V. Matyushov, J. Phys. Chem. A, **106**, (2002), 2146–2157.
27. S. Miertus, E. Scrocco, J. Tomasi, Chem. Phys. **55**, (1981), 117-129.
28. W. T. Ang, *A Beginner's Course in Boundary Element Methods*, Universal Publishers, Boca Raton, 2007.
29. E. Cancés, B. Menucci, J. Tomasi, J. Chem. Phys, **107**, (1997), 3032-3041.
30. C. Amovilli, V. Barone, R. Cammi, E. Cancés, M. Cossi, B. Menucci, C. S. Pomelli, J. Tomasi, Adv. Quantum Chem., **32**, (1998), 227-261.

31. E. L. Coitino, J. Tomasi, R. Cammi, J. Comp. Chem., **16**, (1995), 20-30.
32. R. Cammi, L. Frediani, B. Mennucci, J. Tomasi, K. Ruud, K. V. Mikkelsen. J. Chem. Phys., **117**, (2202), 13-26.
33. F. J. Olivares del Valle, J. Tomasi, Chem. Phys., **150**, (1991), 139-150.
34. M. A. Aguilar, F. J. Olivares del Valle, J. Tomasi, Chem. Phys., **150**, (1991), 151-161.
35. F. J. Olivaeres del Valle, M. A. Aguilar, J. Comp. Chem., **13**, (1992), 115-134.
36. R. Cammi, J. Chem. Phys., **131**, (2009), 164104/1-164104/14.
37. M. Caricato, G. Scalmani, G. W. Trucks, M. J. Frisch, J. Phys. Chem. Lett.,1, (2010), 2369-2373.
38. M. Caricato, G. Scalmani, M. J. Frisch, J. Chem. Phys., **134**, (2011), 244113/1-244113/10.
39. M. Caricato, G. Scalmani, J. Chem. Theo. Comp.,**7**, (2011), 4012-4018.
40. R. Cammi, Adv. Quantum Chem., **64**, (2012), 1-29.
41. C. S. Pomelli, *Cavity Surfaces and their Discretization*, in: Continuum Solvation Models in Chemical Physics: From Theory to Applications, Ed.: B. Mennucci, R. Cammi, John Wiley & Sons Ltd, Hoboken, 2007, pp 49−63.
42. A. Bondi, J. Chem. Phys., **68**, (1964), 441-451.
43. M. L. Connolly, J. Appl. Crystallogr., **16**, (1983), 548-558.
44. J. L. Pascual-Ahuir, *GEPOL: Un metodo para calculo de superficies moleculares*, Tesis Doctoral, València, 1988.
45. A. Klamt, G. Schüürmann, J. Chem. Soc., Perkin Trans. 2, (1993), 799-805.
46. A. Schäfer, A. Klamt, D. Sattel, J. C. W. Lohrenz, F. Eckert, Phys. Chem. Chem. Phys., **2**, (2000), 2187-2193.
47. M. Cossi, N. Rega, G. Scalmani, V. Barone, J. Comp. Chem., **24**, (2003), 669-681.
48. E. Cancès, B. Mennucci, J. Chem. Phys., **115**, (2001), 6130-6135.
49. D. M. Chipman, Theor. Chem. Acc., **107**, (2002), 80-89.
50. D. M. Chipman, J. Chem. Phys., **106**, (1997), 10194-10206.
51. M. Wen, J. Jiang, Z. X. Wang, C. Wu, Theor. Chem. Acc, **133**, (2014), 1471/1-1471/11.
52. D. W. Tondo, J. R. Pliego, J. Phys. Chem. A 109, (2005), 507-511.
53. A. Klamt, *COSMO-RS From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, Elsevier, Amsterdam, 2005.
54. A. Klamt, F. Eckert, W. Arlt, Annu. Rev. Chem. Biomol. Eng., **1**, 2010, 101−122.
55. A. Klamt, WIREs Comput, Mol. Sci., **1**, 2011, 699-709.
56. A. Klamt, V. Jonas, T. Bürger, J. C. W. Lohrenz, J. Phys. Chem. A, **102**, (1998), 5074-5085.
57. E. A. Guggenheim, *Mixtures*, Oxford University Press, Oxford, 1952.
58. G. M. Kontogeorgis, G. K. Folas, *Thermodynamic Models for Industrial Applications*, Wiley, Chichester 2010.
59. J. Gmehling, B. Kolbe, M. Kleiber, J. Rarey, *Chemical Thermodynamics for Process Simulation*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2012.
60. C. Panayiotou, Ind. Eng. Chem. Res., **42**, (2003), 1495-1507.
61. A. Klamt, G. J. P. Krooshof, R. Taylor, AIChE J., **48**, (2002), 2332-2349.
62. R. Franke, J. Friedrich, Chem. Phys., **356**, (2009), 110-120.
63. R. Franke, B. Hannebauer, Phys. Chem. Chem. Phys., **13**, (2011), 21344-21350.
64. R. Franke, B. Hannebauer, S. Jung, Fluid Phase Equilib., **340**, (2013), 11-14.

Modeling Biomolecular Solvation Effects by Hybrid QM/MM Methods

Magnus Schwörer and Gerald Mathias

Lehrstuhl für BioMolekulare Optik,
Ludwig–Maximilians Universität München, Oettingenstr. 67, 80538 München, Germany

E-mail: gerald.mathias@physik.uni-muenchen.de

The combination of quantum mechanical (QM) methods with molecular mechanics (MM) force fields in QM/MM hybrid simulations is a powerful approach to compute molecular properties or chemical reactions in the condensed phase, e.g. in aqueous solution or in large biomolecular complexes. Many flavors of QM/MM have spawned during the last four decades of development and successful application. We will review the fundamental ideas of QM/MM and important considerations for their practical deployment. Assessment of the accuracy of QM/MM methods and strategies for their advancement are discussed.

1 Introduction

Nearly 40 years ago in 1976 the QM/MM concept for molecular simulations — that is to employ quantum mechanical (QM) methods to a region of interest while the surrounding environment is treated by molecular mechanics (MM) force fields — was introduced by Arieh Warshel and Michael Levitt in their seminal study on the carbonium ion in lyzozyme.[1] This first QM/MM publication already provided the key ingredients that are still used in hybrid QM/MM simulations today although methods and computational details have greatly evolved. The huge impact of QM/MM on molecular science spanning the fields of physics, chemistry, and biology has lately been acknowledged by awarding the 2013 Nobel price in chemistry to these two authors along with Martin Karplus.[2]

Basis for most QM/MM implementations is the partitioning of the Hamiltonian[1]

$$H = H_{\mathrm{MM}} + H_{\mathrm{QM}} + H_{\mathrm{QM/MM}} \tag{1}$$

of the system under investigation into interactions H_{MM} within the MM subsystem, interactions H_{QM} within the QM subsystem, and a coupling term $H_{\mathrm{QM/MM}}$ modeling the interaction between the two subsystems. Many flavors of QM/MM methods have been derived from (1) altering the choice of the MM force field or the particular QM method, such as various semi–empirical (SE) approaches, Hartree–Fock (HF) and post–HF methods, as well as density functional theory (DFT). Furthermore, the details of the coupling term $H_{\mathrm{QM/MM}}$ have evolved in accuracy as well as in efficiency. Accordingly, this zoo of methods spans a large range of accessible sizes of QM subsystems, accuracies, and time scales in molecular simulations, which one always has to keep in mind when conducting, evaluating and comparing QM/MM studies. Overviews on QM/MM are provided by many excellent reviews.[3–8]

The excellence of QM/MM methods is to account for the effects of the solvent environment on the molecular properties accessible by an accurate QM electronic structure computation. In the QM/MM calculation the solvent is treated by an efficient MM force

field such that the total costs are far less than those of a full QM condensed–phase simulation or a QM computation for a large cluster model. In this context we first think of the solvent as an aqueous solution consisting of water molecules and possibly ions which thermally sample many configurations around a QM solute. Equally well, we can conceive compounds of biomolecules and substrates within biomolecules as solutes in complex and specialized solvent environments, for example a reactive site solvated within an enzyme. Here, the environment is typically well defined by the tertiary structure of the enzymatic protein, which provides specific electrostatic and steric interactions with the QM fragment, e.g. by polar and charged amino acid side chains. Still, at ambient conditions such solvent environments for the reactive sites may show a certain degree of fluctuations for example in the hydrogen bonding network to and around the substrate or the orientation of side chains.

We will review the key ideas, the necessary ingredients and computational obstacles to model such solute–solvent systems within the QM/MM framework. After a short review of the basics of QM/MM coupling methods we will discuss examples on how one can improve their efficiency and accuracy both on the level of the force field and the coupling scheme.

2 QM/MM Interactions

Fundamental for all QM/MM methods is the specification of the coupling term

$$H_{\mathrm{QM/MM}} = H^{\mathrm{elec}}_{\mathrm{QM/MM}} + H^{\mathrm{vdW}}_{\mathrm{QM/MM}} + H^{\mathrm{cov}}_{\mathrm{QM/MM}} \tag{2}$$

entering Eq. (1), which covers electrostatic and van der Waals (vdW) interactions between the QM and MM fragments, as well as the treatment of possible covalent bonds between the two fragments.

2.1 Breaking Covalent Bonds

For the latter $H^{\mathrm{cov}}_{\mathrm{QM/MM}}$, the challenge is to model the broken and thus unsaturated covalent bond without spuriously altering the electronic structure of the QM fragment in the region of interest. Here, a simple strategy is to saturate the broken bond by an additional link atom, typically a hydrogen atom, which seems justified for single sigma–type bonds, such as homopolar C–C bonds.[9] The particular placement of the link atom and possible correction terms help to minimize the distortion of the mechanics and the electronic structure of the link region.[10,5] An alternative route is to prepare special QM atoms at the boundary between QM and MM fragment, such that their electrons still model the formally broken bond. In the frozen orbital method this is achieved by fixing the coefficients of certain hybrid orbitals to mimic a covalent bond to the MM fragment.[5] A related idea is to construct special pseudopotentials for such boundary atoms in DFT calculations, which the cover the electrons in frozen orbitals in addition to core electrons, such that the link atoms have a reduced valence.[11] Comprehensive and complete overviews on the issue of breaking covalent bonds can be found in a recent review article Ref. 5.

2.2 Van der Waals Interactions

The vdW interactions $H_{\text{QM/MM}}^{\text{vdW}}$ have to cover the Pauli repulsion and the dispersive interactions between the electrons of the two fragments. These interactions are modeled at the force field level and one typically adopts the functional form and the parametrization of the MM force field. Note that particularly the repulsive terms are crucial, since they essentially determine the distances between QM and MM atoms, e.g. in hydrogen bonds between the two fragments. Further below in Sec. 4.1 we will discuss how a QM/MM description can be improved by proper parametrization of inter–fragment vdW interactions.

2.3 Electrostatic Interactions

The computationally most demanding part of the QM/MM interactions are the electrostatic interactions

$$H_{\text{QM/MM}}^{\text{elec}} = \int \mathrm{d}r^3 \varrho_{\text{el}}(\mathbf{r})\Phi^{\text{MM}}(\mathbf{r}) + \sum_{\mu=1}^{N^{\text{QM}}} Z_\mu \Phi^{\text{MM}}(\mathbf{R}_\mu) \tag{3}$$

of the N^{QM} core charges Z_μ of the QM atoms μ at the positions $\mathbf{R}_\mu$ and their electron charge density ϱ_{el} with the external potential Φ^{MM} generated within the QM region by the MM atoms. Most simple so–called *subtractive* or *additive mechanical embedding* QM/MM schemes do not include the electrostatic perturbation by Φ^{MM} in the QM Hamiltonian. For most solvation effects such simple models are much too crude and *additive electrostatic embedding* schemes have to be used, in which the MM fragment fully polarizes the QM electronic structure, as in the original work of Warshel and Levitt.

The demanding part of Eq. (3) is to calculate the energy of the electrons in the external potential Φ^{MM}. Employing the linear combination of atomic orbitals (LCAO) approach in the QM calculation and a simple Coulomb model for the MM atoms these interactions yield the coupling terms

$$h_{\mu,k}^{\mu',k'} = -\sum_{i=1}^{N^{\text{MM}}} \frac{e\, q_i}{4\pi\varepsilon_0} \int \mathrm{d}r^3 \frac{\varphi_{\mu,k}^*(\mathbf{r})\varphi_{\mu',k'}(\mathbf{r})}{|\mathbf{r}-\mathbf{R}_i|} \tag{4}$$

between the atomic orbitals $\varphi_{\mu,k}$ and $\varphi_{\mu',k'}$ on QM atoms μ and μ' and the partial charges q_i of the MM atoms located at positions $\mathbf{R}_i$. The number of coupling terms scales quadratically with the number of atomic orbitals and linearly in the number of MM atoms. Similar three–center integrals have to be evaluated for the interactions of the electrons with the QM cores, but, since the MM atoms typically outnumber the QM atoms by a few orders of magnitude, their number is much larger for QM/MM interactions. Note that for SE or tight–binding DFT calculations these integrals do not have to be solved explicitly but one can employ the same approximations for the MM charges as one uses for the QM cores.[9,12,13]

If, instead, the electron density can be represented on a spatial grid such as in plane–wave basis or hybrid basis DFT implementations,[14,15] the electron–MM interactions

$$\int \mathrm{d}r^3 \varrho_{\text{el}}(\mathbf{r})\Phi^{\text{MM}}(\mathbf{r}) = \sum_{\gamma=1}^{N_{\text{grid}}} \varrho_{\text{el}}(\mathbf{r}_\gamma)\Phi^{\text{MM}}(\mathbf{r}_\gamma) \tag{5}$$

reduce to a simple sum over the N_{grid} grid points γ at $\mathbf{r}_\gamma$. Here, the computational task is to map Φ^{MM} on the grid used in the QM calculation, which can easily span a few million grid points.

2.4 Computational Aspects

For extended simulations both variants of QM/MM electrostatic interactions (4) and (5) are prohibitively expensive in a typical QM/MM simulation setup if employed naively. Therefore, interactions are often truncated after a certain cutoff distance R_c between QM and MM atoms to limit the number of QM/MM interactions. Plain truncation schemes, however, are notorious for introducing artifacts, as has been shown in many MM–MD studies,[16,17] which limits the use of plain truncation to more or less static calculations. Moreover, truncation around a reactive site within a protein may just include or exclude important interactions with charged amino acid side chains and, thus, the results of the calculations may strongly depend on R_c.

Multiple-scale strategies increase the accuracy of the QM/MM electrostatics computation as compared to straight truncation schemes by treating more distant regions by computationally efficient approximations. Some schemes expand the charge density of the QM fragment in multipole moments for its interaction with distant MM atoms[11] or use electrostatic potential derived charges (ESP) of the QM fragment for interactions with MM charges.[18] Fast multipole methods, which group distant atoms into hierarchically nested clusters and use multipole and local Taylor expansions to compute their collective interactions with the QM fragment enable even more efficient QM/MM computations.[10,19,20] Electrostatic interactions beyond the explicitly treated region can be furthermore complemented by interactions with a dielectric continuum, which exerts a reaction field (RF) on the QM region and shields long–range interactions.[21–31] For instance, excitation energies and magnetic shielding constants of diazines calculated by such a multilayer QM/MM/RF model yielded a very good agreement with a large explicit solvent treatment.[27] For many applications the size of the explicit solvent region can be decreased by such a multilayer method,[29] and the errors entailed by truncation of the electrostatic interactions are suppressed.[24,32] Alternatively the QM/MM calculation can be embedded in a fully periodic scheme for the MM fragment based on Ewald summation,[11,33] which also avoids truncation effects.

Another multiple-scale approach for the reduction of the computational effort is to map the MM potential Φ^{MM} onto coarser grids for more distant MM atoms, which can then be interpolated to yield a very good approximation of Φ^{MM} on the finest grid eventually needed for the DFT calculation.[33]

In QM/MM calculations Φ^{MM} has to be imported to the QM calculation but also the forces on the MM atoms have to be calculated after the QM self–consistent field (SCF) iterations have converged. All modern approximative schemes for QM/MM electrostatics calculations use fully Hamiltonian forces, i.e. based on derivatives of the approximate MM potential used in the QM calculation with respect to the MM positions $\mathbf{R}_i$.[11,33,19] Note that in Car–Parrinello (CP) simulations,[34] which do not rely on an SCF procedure but integrate an extended Lagrangian scheme, the import of Φ^{MM} and the export of forces can be done simultaneously.[11] Still, QM/MM CP–MD simulations may require in total more such import and export operations due to the much smaller time step required for CP–MD

as compared to SCF–based Born–Oppenheimer–MD, which may be an efficiency issue that has to be considered when choosing the method of simulation.

3 Determining the QM/MM Setup

Setting up QM/MM simulations is in many cases a painful iterative approach due to the complexity of the method and the peculiarities of the system under investigation. Especially if the project will involve significant amounts of computer time it is worthwhile to invest considerable effort in the proper choice of the setup.

3.1 The QM Method

The requirements for the QM level of theory in a QM/MM study are mostly the same as for pure QM studies. The applied method must be capable to correctly model the effects and properties that are targeted for the system under investigation at a reasonable computational cost. Mostly it is advisable to first study a possibly simpler pure QM system as a reference in order to calibrate the accuracy of the method. Moreover, such preparatory calculations help to quantify the effects of solvation in the following main QM/MM study.

Principle choices are SE methods such as MNDO[35] and AM1,[36] or tight–binding DFT methods[13] which require a comparable computational cost. SE methods are fast and can yield high throughput for a wide selection of organic molecules. Still, it is essential to check if the underlying SE parametrization is suited for the system under investigation, and, if necessary to reparametrize the method. In contrast, HF methods are parameter free, but computationally much more demanding. HF is the canonical precursor for post–HF methods, which are required to address properties of electronic exited states, such as optical spectra. Post–HF methods require careful calibrations to minimize the computational cost for the desired accuracy. For the electronic ground state, Kohn–Sham DFT[37] yields a good compromise between accuracy and efficiency and is therefore widely used for QM/MM–MD investigations. DFT requires the choice of the exchange and correlation functional, possibly pseudopotentials, and, similar to HF, a suitable basis set.

A second important criterion is whether the QM method supplies fast analytical gradients necessary for molecular dynamics (MD) simulations, which have to be conducted whenever one aims at extended sampling for free energy calculations or time correlation functions, e.g. for the calculation of vibrational spectra.[38,39]

3.2 The QM/MM Interface Software

Having determined the QM method of choice, the next step is to choose a suitable interface software for QM/MM. Modular approaches like the ChemShell,[40,41] QMMM,[42] or QuanPol[43] packages offer interfaces to many QM methods implemented in various QM programs and to several MM force fields and MD programs. A uniform interface software between the respective programs deals with the import and export of electrostatic potentials and forces on the atoms.

On the other hand, specialized interface implementations coupling grid–based DFT methods very efficiently to MM–MD programs[10,11,33,19,20] or implementing SE Hamiltonians in MM–MD programs[36,44–46,12] have the advantage of fast data exchange and optimized import and export routines as we have discussed for the electrostatic coupling

above. These aspects are of particular importance, whenever the MM fragment and the interface part of the calculation consume a significant portion of the available computational resources.

3.3 The MM Force Field

Choosing a specialized interface often implies the selection of a certain MM–MD package. Many such packages are shipped with a native MM force field, for example the GROMOS force field[47] for the CPMD/GROMOS coupling[11] or AMBER[48] and CHARMM[49] for their native SE coupling.[44,45,12] However, due to the functional similarity of today's popular MM force fields for biomolecular simulations a conversion of input files between different program formats is often feasible at moderate effort. This can be useful if one already has set up the pure MM system with a particular force field and verified the stability of the simulation, particularly for the region of interest, in extended MM–MD simulations. Whether certain MM force fields yield better QM/MM results than others is rarely investigated[50] and one may only guess if a more or a less polar parametrization is preferential for the system of interest. Special care is required when the force field employs unified atom models[51,52] (e.g. use a single atom for a CH_2 group) to replicate the hydrogens, if such atoms are included in the QM region.[11]

Stable MM–MD simulations and initial equilibration runs are a neccessary prerequisite if MD is the sampling method of choice for the QM/MM study. For the MM equilibration one may have to parametrize prostetic groups, which are not included in the MM force field, although they will be treated by QM methods in the productive phase of the simulations. Alternatively, one can describe the prostetic group by less expensive lower level QM methods such as SE during the equilibration phase.

3.4 The QM Fragment

Finally, one has to properly partition the system into the QM and the MM region, which is a most crucial step for the balance of efficiency and accuracy of the simulation. Here, the property or the chemical reaction under investigation typically provides a minimal choice for the QM region, e.g. the set of all reactants. For prostetic groups or other parts of macromolecules within the QM region the next step is to select covalent bonds between the QM and MM region that can be cut as discussed above in Sec. 2.1. Here, the distortion induced by this boundary between QM and MM region should be as small as possible for the chemistry described in the QM region, see Ref. 5 for a complete review.

For enhanced accuracy one may extend the minimal QM region by further fragments, e.g. by polar and charged amino acid side chains around the region of interest, thereby treating important mutual polarization effects on the QM level. Similarly, also interactions with the solvent may be of utmost importance for the property under investigation. It is therefore intriguing to reduce possible errors introduced by the force field by extending the QM region to cover solvent molecules in the inner solvation shell(s) thereby moving the QM/MM boundary further away from the solute.[9] Treating parts of the solvent by QM raises the problem of how to treat MM solvent molecules that diffuse into the QM region and, vice versa, QM solvent molecules that leave it during a MD simulation. One can either restrain the QM and MM solvent molecules to stay in the respective QM (MM) regions,

or construct an *adaptive* QM/MM scheme, whose main challenge is to avoid huge energy and force discontinuities while still being computationally feasible; for a recent overview see Ref. 53.

Such extended QM regions entail, of course, enhanced computational costs, which may be prohibitively large if expensive *ab inito* methods are used for the QM region and long QM/MM–MD trajectories are needed for proper sampling.

4 Probing and Improving the Accuracy of QM/MM Calculations

The previous overview of the aspects of proper QM/MM setup revealed that the required choices may influence the results of the simulation and that one inevitably needs to cross–check to what extent the results and the conclusions depend on the QM/MM setup. Here, one important test is to vary the size of the QM region. For example, to probe ionic solvation in water with QM/MM one needs to include up to two solvent shells in the QM fragment to obtain consistent results for the solvation structure.[54,55] Similarly, a single QM glycine in MM water does not retain its native zwitterionic state, but was found to be stabilized by including a first water shell in the QM fragment.[56] Meier et al. show that they can alter the hydrogen bonding pattern of the AppA BLUF protein around its FMN cofactor depending on the amino acid side chains included in the QM fragment.[50] An even stronger qualitative dependence on the size of the QM fragment was found for the phosphate monoester hydrolysis in alkaline phosphatase.[57] Here, the inclusion of embedded zinc ions and their coordinating amino acid side chains changed the predicted mechanism from an associative reaction path to a dissociative reaction path.

All these examples seem to indicate that larger QM regions should be preferred over minimal setups. A different interpretation is that the force fields employed in QM/MM simulations model the interactions of the MM fragment with the QM fragment only at insufficient quality for many properties and systems of interest. Therefore, an alternative approach to increasing the size of the QM region is to improve the accuracy of the force field as well as the accuracy the mutual interactions of QM and MM fragment to obtain results consistent with those of larger QM regions.

4.1 Adapting van der Waals Parameters

In MM force fields, both the vdW intermolecular interactions and the electrostatic interactions are specifically adjusted according to the respective parametrization strategy and depend, of course, on the functional form of the employed interaction potentials. Their balance determines the distance and strengths of hydrogen bonds or the structure of radial distribution functions. Because the point charges of the MM force field yield rather strong short–range Coulombic forces, they are typically counterbalanced by 12–6 Lennard–Jones (LJ) potentials which feature stiff repulsion. However, the electrostatic interactions between MM and QM atoms described by $H^{\mathrm{elec}}_{\mathrm{QM/MM}}$ in Eq. (2) are qualitatively different to MM–MM interactions because within the QM fragment the singular partial charges of the force field are replaced by a much smoother electron charge density shielding the core charges. Therefore, one also expects the short–range electrostatic QM/MM interactions to be weaker than pure MM electrostatic interactions.

Sticking to the LJ potential and its MM parametrization also for the QM/MM vdW interactions, as most QM/MM methods do,[5] may exaggerate the vdW repulsion. Simply adapting the force fields' vdW parameters was reported to yield erroneous binding energies,[58] geometries[59] and solvation structures.[60]

The choice of vdW parameters plays a decisive role for the solvation structures around the QM fragment as measured by the radial distribution function, and, therefore, for the correct polarization of a QM solute. In contrast, thermodynamic properties like reduction potentials or potentials of mean force are less sensitive to the choice of vdW parameters, which was attributed to a cancellation of errors.[60]

VdW parameters for QM/MM settings can be optimized to reproduce the hydrogen bonding energies, distances or geometries for selected representative molecular clusters in the gas phase,[61,62,58,63,60,64,59] for which full QM calculations at the *ab initio* level may serve as references. The reference should, of course, correctly model hydrogen bond structures, which may require that also dispersion interactions are reasonably represented by the QM method. Such optimized parameters can then also predict hydrogen bonding for a test set of larger molecular complexes, which were not part of the parametrization set, at high accuracy.[62] The parameters optimized for a specific QM/MM method and MM force field depend on the form of the respective QM/MM interaction Hamiltonian,[63] and are in general not transferable between different QM/MM methods, different basis sets,[64] or excitation states.[65]

Also, the transferability of optimal vdW parameters obtained for gas phase clusters to the bulk is questionable. Tu and Laaksonen found that the optimal vdW parameters for a QM/MM water dimer in vacuum rendered the solute–solvent interactions in the condensed phase too weak, resulting in an underestimation of the QM water molecule's induced dipole moment,[66] i.e. in incorrect polarization. Similarly, the peak positions of the radial distribution function were shifted to larger distances, as compared to experimental and pure MM results. The standard MM force field parameters, in contrast, entailed too strong QM/MM interactions.

Alternative parametrization strategies rely on thermodynamic properties of solvated QM molecules,[67,68] but are computationally more expensive than gas phase cluster calculations. Here, a computationally efficient strategy like the "averaged solvent electrostatic potential" (ASEP) approach[69,70] proved to be valuable to iteratively optimize the vdW parameters to bulk properties.[71] Employing this technique, QM/MM vdW parameters were successfully fitted to experimental hydration enthalpies and radial distribution functions. In the subsequent simulations they notably improved the QM/MM modeling of ammonia ionization.[72] For small solvated molecules such efficient mean–field based QM/MM calculations can yield reliable structural properties.

The LJ potential widely used in QM/MM applications has deficiencies in modeling short–range intermolecular forces.[73] Improved potential energy curves for small molecular have been found with a Buckingham– or pseudopotential–like modeling of the short–range repulsion.[73,58] Furthermore, the LJ potential only empirically accounts for Pauli repulsion and does not act on the QM electron density. The missing Pauli repulsion may substantially affect the electron density and thus, e.g. the magnetic properties of QM molecules.[74] Also here, augmenting MM atoms with effective repulsive (pseudo)potentials entering the QM Hamiltonian may improve the modeling.

In summary, the choice of vdW parameters can significantly alter the solvation struc-

tures of molecules, and reparametrization of vdW interactions needs to be considered for an optimal modeling of QM/MM solvation.

4.2 Polarization of the MM Fragment

The above discussion showed that in particular at short ranges the description of the vdW and electrostatic interactions mutually depend on each other. So, a next important step is to improve the electrostatic description of the MM environment around the QM fragment.

The severe shortcoming of today's popular biomolecular MM force fields is the lack of explicit electrostatic polarization terms. Polarization is only included in a mean–field fashion, i.e. MM force fields use the same partial charges for an amino acid side chain independent of its local environment. Moreover, these partial charges are not necessarily parametrized to reproduce a mean polarization but also depend on other quantities such as free energies of solvation. On the one hand this mean–field approach allows to tackle larger and larger systems and to reach longer simulation time scales with the increasing computational power. On the other hand it is questionable if the current unpolarizable MM force fields are accurate enough to be predictive on these length and time scales.[75] It is widely accepted that, in order to improve force field accuracy, electronic polarization effects should be included.[76–79]

4.2.1 Polarized MM Force Fields

A first step to improve the electrostatic description of the MM environment is to locally polarize it. Here, the key idea is to adapt the mean–field partial charges of the amino acid side chains around the QM fragment to the specific local environment. For example, using standard force field parameters in MM–MD simulations of the membrane protein bacteriorhodopsin (BR) leads to collapse of the region around the retinal chromophore,[80] which is, however, well defined and stable according to the crystal structure.[81] After adapting the partial charges of the MM force field around the chromophore by DFT/MM calculations to the local electrostatic fields within the chromophore binding pocket the resulting *polarized* force field turned out to stabilize the structure. Moreover, the infrared (IR) spectrum of BR's chromophore calculated by DFT/MM in this polarized environment was accurate enough to distinguish the native structure from non–native protein conformations by comparison with experimental results.[82] The same strategy of polarizing the amino acid side chains in the MM force field was applied in a DFT/MM IR study of the FMN cofactor in various crystal structures of the AppA BLUF protein.[83] Here, the improved quality of the polarized MM description enhanced the accuracy of the computed IR spectra such that the quality of the crystal structures could be judged.[83]

These findings show the importance of a correctly polarized environment for properties sensitive to solvation and electrostatic interactions, such as IR spectra. The development of a polarized force field can be technically quite involved and pays off only if in turn one can use a minimal QM fragment for high–level QM methods. Moreover, a polarized force field is specific to a certain protein in a well determined conformation, which makes it suitable only for comparatively rigid systems, like the retinal binding pocket in BR. Neither are polarized force fields transferable nor do they enable unbiased conformational sampling.

4.2.2 Polarizable MM Force Fields

Polarizable MM force fields (PMM),[76–79] in contrast, instantaneously model polarization effects and account for the fluctuating environment. There are various approaches to model the polarization of the PMM force field. The fluctuating charge (FC) approach[78,84] redistributes the partial charges depending on the external electric field. For planar molecules this approach only covers the components of the polarization within the molecular plane. Drude oscillators (DO)[78,85] model atomic induced dipoles by attaching an additional partial charge on a stiff short harmonic spring to each polarizable atom. The induced dipole (ID) approach[78] explicitly introduces polarizable dipoles, whose strengths and directions depend linearly on the electric field. A short–range damping scheme is usually employed to avoid overpolarization at short distances.[86] This approach has been generalized to inducible multipoles by the AMOEBA force field.[87]

Polarization models are not pairwise additive like Coulomb or LJ interactions. Therefore, the polarization state of a given atomic configuration of a system needs so be determined self–consistently, i.e. the strengths of the IDs, the positions of the DOs, or the partial charges of the FC groups, respectively. Alternative to this SCF approach, the polarizable degrees of freedom can be treated as independent dynamical variables by an extended Lagrangian approach, similar to CP–MD,[78] which, however, requires a smaller integration time step. Therefore, polarizable force fields are computationally more expensive. Furthermore, DO or ID models have an increased number of interaction partners and ID models require the evaluation of more complex interaction functions. Furthermore, the additional complexity of PMM force fields requires suitable parametrization strategies.[84,85,88,87,89]

For water, polarizable models can render correct gas–phase electrostatic properties as well as a wide range of bulk water properties.[78,89] Their transferability to other environments like protein–water mixtures should be superior to unpolarizable MM models by construction. In particular for ionic solutions, in which water molecules are exposed to large local electric fields, polarizable models should excel mean–field models.

4.2.3 MM Polarization in QM/MM

QM/MM schemes are expected to profit from the enhanced accuracy of PMM force fields in general, and of polarizable water models in particular for aqueous solutions, but the QM method has first to be properly interfaced with the polarizable force field. To include polarizable degrees of freedom into an existing QM/MM method, their electrostatic interaction energy with the QM fragment $H^{\mathrm{pol}}_{\mathrm{QM/MM}}$ has to be included in the hybrid Hamiltonian (2).[1] For PMM force fields based on FC or DO, i.e. polarization modeled by charges, this is straightforward, as the functional form of the interactions is identical to those already implemented for QM/MM interactions.[90–93,30] Interfacing polarizable ID with a QM method requires more effort[1,9,19,20,94] as qualitatively different QM/MM interaction terms to Eq. (4) have to be considered.

For QM/PMM calculations the SCF conditions must be met not only for the individual QM and PMM fragments but also for their mutual interaction. Thompson and Schenter presented a corresponding iteration scheme, which alternately iterates the QM and PMM subsystems to full convergence until global convergence is reached.[95,96] The efficiency of solving the QM/PMM SCF can be enhanced by alternating between QM and PMM

fragment after each iteration step[65,97] or by a convergence–driven switching between the QM and PMM iterations.[19]

Although Warshel and Levitt included explicit MM polarization in their hybrid scheme and stressed the importance of polarization effects,[1] the majority of QM/MM methods and applications employ unpolarizable MM force fields.[5] With the availability and growing hands–on experience with PMM force fields over the past decade,[78] they have also reemerged in QM/PMM studies, in particular in implementations which supply atomic forces for geometry optimization or MD simulations.[98,95,96,91,99,92,65,93,100,101,29,30,19,20,94]

QM/PMM studies reported substantial solute–solvent energy contributions by MM solvent polarizability,[1,102,103] and more accurate solvation effects for the excitation energies of rhodopsins[104,105] and uracil,[106] while the effect of MM polarization on optimized geometries is less pronounced.[9,94]

The correct modeling of polarization effects in the MM surrounding is especially important for charged QM fragments.[107] MM polarization was found to be crucial when one calculates the proton affinities and deprotonation enthalpies of alcohols,[9] or solvation of fluoride[99] and chloride ions,[93] or cation/benzene complexes.[108] Relative solvation free energies of Mg^{2+} and Zn^{2+} ions computed from MD simulations using a combination of DFT and a PMM water model yielded very good agreement with experiment.[55]

These examples demonstrate that a proper inclusion of polarizability in the MM fragment improves the accuracy of QM/MM results. Since efficient QM/PMM–MD implementations have become available,[19,20] one can proceed to benchmark such methods by long trajectories. Concomitantly, high–level QM–MD simulations of bulk systems have become feasible[109,110,88,111] which render abundant reference data to evaluate the accuracy of QM/(P)MM methods. Furthermore, these data can guide the parametrization of PMM force fields and QM/PMM interfaces.[112,88]

5 Summary

With the development of QM/(P)MM methods molecular properties and chemical reactions in the condensed phase can be computed at high accuracy, while the costs are mainly dictated by the number of atoms included in the QM fragment and the chosen QM method. Selecting the proper method for the system of interest requires considerations concerning the level of QM theory, the QM/MM interface, the MM force field, and the size of the QM fragment. Adapting the vdW interactions between QM and MM fragment as well as to take the electronic polarization of the MM fragment into account can greatly improve the accuracy of QM/MM methods.

The effort to familiarize with the various aspects of QM/MM is well invested, since they open a large field of possible applications, many of which currently are on the verge of emerging.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (SFB 749/C4) and by the Bavarian Competence Network for Technical and Scientific High Performance Computing (KONWHIHR-III).

References

1. A. Warshel and M. Levitt, *Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme*, J. Mol. Biol., **103**, 227–249, 1976.
2. M. Levitt, *Birth and future of multiscale modeling for macromolecular systems (Nobel Lecture).*, Angew. Chem. Int. Ed. Engl., **53**, 10006–18, 2014.
3. J. Gao and D. G. Truhlar, *Quantum mechanical methods for enzyme kinetics.*, Annu. Rev. Phys. Chem., **53**, 467–505, 2002.
4. H. Lin and D. G. Truhlar, *QM/MM: what have we learned, where are we, and where do we go from here?*, Theoret. Chem. Acc., **117**, 185–199, 2006.
5. H. M. Senn and W. Thiel, *QM/MM Methods for Biomolecular Systems*, Angew. Chem. Int. Ed., **48**, 1198–1229, 2009.
6. F. H. Wallrapp and V. Guallar, *Mixed quantum mechanics and molecular mechanics methods: looking inside proteins*, Comput. Mol. Sci., **1**, 315–322, 2011.
7. M. W. van der Kamp and A. J. Mulholland, *Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology.*, Biochemistry, **52**, 2708–28, 2013.
8. B. Mennucci, *Modeling environment effects on spectroscopies through QM/classical models.*, Phys. Chem. Chem. Phys., **15**, 6583–94, 2013.
9. D. Bakowies and W. Thiel, *Hybrid models for combined quantum mechanical and molecular mechanical approaches*, J. Phys. Chem., **100**, 10580–10594, 1996.
10. M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello, *A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields*, J. Chem. Phys., **110**, 10452–10467, 1999.
11. A. Laio, J. VandeVondele, and U. Rothlisberger, *A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations*, J. Chem. Phys., **116**, 6941–6947, 2002.
12. Q. Cui, M. Elstner, E. Kaxiras, T. Frauenheim, and M. Karplus, *A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method*, J. Phys. Chem. B, **105**, no. 2, 569–585, 2001.
13. M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, *Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties*, Phys. Rev. B, **58**, 7260–7268, 1998.
14. © IBM Corp 1990–2008 and MPI für Festkörperforschung Stuttgart 1997–2001, `www.cpmd.org`, *CPMD: Car–Parinello Molecular Dynamics, version 3.17.1*.
15. J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, *Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach*, Comput. Phys. Commun., **167**, 103 – 128, 2005.
16. P. H.. Hünenberger and W. F. van Gunsteren, *Alternative schemes for the inclusion of a reaction-field correction into molecular dynamics simulations: Influence on the simulated energetic, structural, and dielectric properties of liquid water*, J. Chem. Phys., **108**, 6117–6134, 1998.
17. G. Mathias, B. Egwolf, M. Nonella, and P. Tavan, *A fast multipole method combined with a reaction field for long-range electrostatics in molecular dynamics simulations: The effects of truncation on the properties of water*, J. Chem. Phys., **118**, 10847–10860, 2003.

18. A. Laio, J. VandeVondele, and U. Rothlisberger, *D-RESP: Dynamically Generated Electrostatic Potential Derived Charges from Quantum Mechanics/Molecular Mechanics Simulations*, J. Phys. Chem. B, **106**, 7300–7307, 2002.
19. M. Schwörer, B. Breitenfeld, P. Tröster, S. Bauer, K. Lorenzen, P. Tavan, and G. Mathias, *Coupling density functional theory to polarizable force fields for efficient and accurate Hamiltonian molecular dynamics simulations.*, J. Chem. Phys., **138**, 244103, 2013.
20. M. Schwörer, K. Lorenzen, G. Mathias, and P. Tavan, *Utilizing Fast Multipole Expansions for Efficient and Accurate Quantum-Classical Molecular Dynamics Simulations*, J. Chem. Phys., **142** (11), xxx, 2015.
21. Q. Cui, *Combining implicit solvation models with hybrid quantum mechanical/molecular mechanical methods: A critical test with glycine*, J. Chem. Phys., **117**, 4720, 2002.
22. M. Nonella, G. Mathias, and P. Tavan, *Infrared Spectrum of* p*-Benzoquinone in Water Obtained from a QM/MM Hybrid Molecular Dynamics Simulation*, J. Phys. Chem. A, **107**, 8638–8647, 2003.
23. B. Mennucci and J. M. Martínez, *How to model solvation of peptides? Insights from a quantum mechanical and molecular dynamics study of N-methylacetamide. 2. 15N and 15O nuclear shielding in Acetone*, J. Phys. Chem. B, **109**, 9830–9838, 2005.
24. P. Schaefer, D. Riccardi, and Q. Cui, *Reliable treatment of electrostatics in combined QM/MM simulation of macromolecules.*, J. Chem. Phys., **123**, 014905, 2005.
25. N. Rega, G. Brancato, and V. Barone, *Non-periodic boundary conditions for ab initio molecular dynamics in condensed phase using localized basis functions*, Chem. Phys. Lett., **422**, no. 4-6, 367–371, 2006.
26. A. H. Steindal, K. Ruud, L. Frediani, K. Aidas, and J. Kongsted, *Excitation energies in solution: the fully polarizable QM/MM/PCM method.*, J. Phys. Chem. B, **115**, 3027–37, 2011.
27. V. Manzoni, M. L. Lyra, K. Coutinho, and S. Canuto, *Comparison of polarizable continuum model and quantum mechanics/molecular mechanics solute electronic polarization: study of the optical and magnetic properties of diazines in water.*, J. Chem. Phys., **135**, 144103, 2011.
28. T. Benighaus and W. Thiel, *Long-Range Electrostatic Effects in QM/MM Studies of Enzymatic Reactions: Application of the Solvated Macromolecule Boundary Potential*, J. Chem. Theory Comput., **7**, 238–249, 2011.
29. F. Lipparini, C. Cappelli, G. Scalmani, N. De Mitri, and V. Barone, *Analytical First and Second Derivatives for a Fully Polarizable QM/Classical Hamiltonian*, J. Chem. Theory Comput., **8**, 4270–4278, 2012.
30. E. Boulanger and W. Thiel, *Solvent Boundary Potentials for Hybrid QM/MM Computations Using Classical Drude Oscillators: A Fully Polarizable Model*, J. Chem. Theory Comput., **8**, 4527–4538, 2012.
31. S. Caprasecca, C. Curutchet, and B. Mennucci, *Toward a Unified Modeling of Environment and Bridge-Mediated Contributions to Electronic Energy Transfer: A Fully Polarizable QM/MM/PCM Approach*, J. Chem. Theory Comput., **8**, 4462–4473, 2012.
32. B. Mennucci and J. M. Martínez, *How to model solvation of peptides? Insights from a quantum-mechanical and molecular dynamics study of N-methylacetamide. 1. Geometries, infrared, and ultraviolet spectra in water.*, J. Phys. Chem. B, **109**, 9818–9829, 2005.

33. T. Laino, F. Mohamed, A. Laio, and M. Parrinello, *An efficient real space multigrid QM/MM electrostatic coupling*, J. Chem. Theory Comput., **1**, 1176–1184, 2005.
34. R. Car and M. Parrinello, *Unified Approach for Molecular Dynamics and Density-Functional Theory*, Phys. Rev. Lett., **55**, 2471–2474, 1985.
35. M. J. S. Dewar and W. Thiel, *Ground states of molecules. 38. The MNDO method. Approximations and parameters*, J. Am. Chem. Soc., **99**, 4899–4907, 1977.
36. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model*, J. Am. Chem. Soc., **107**, 3902–3909, 1985.
37. W. Kohn and L. J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects*, Phys. Rev., **140**, no. 4A, A1133–A1138, 1965.
38. G. Mathias and D. Marx, *Structures and spectral signatures of protonated water networks in bacteriorhodopsin*, Proc. Natl. Acad. Sci. USA, **104**, 6980–6985, 2007.
39. M. Baer, G. Mathias, I. W. Kuo, D. J. Tobias, C. J.Mundy, and D. Marx, *Spectral Signatures of the Pentagonal Water Cluster in Bacteriorhodopsin*, ChemPhysChem, **9**, 2703–2707, 2008.
40. P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjø voll, A. Fahmi, A. Schäfer, and C. Lennartz, *QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis*, J. Mol. Struct.: THEOCHEM, **632**, no. 1–3, 1–28, 2003.
41. "Chemshell, a computational chemistry shell, see `www.chemshell.org`".
42. H. Lin, Y. Zhang, S. Pezeshki, and D. G. Truhlar, "Qmmm 1.3.8, see `http://comp.chem.umn.edu/qmmm`".
43. N. M. Thellamurege, D. Si, F. Cui, H. Zhu, R. Lai, and H. Li, *QuanPol: a full spectrum and seamless QM/MM program.*, J. Comput. Chem., **34**, 2816–33, 2013.
44. M. J. S. Dewar and W. Thiel, *A semiempirical model for the two-center repulsion integrals in the NDDO approximation*, Theoret. Chim. Acta, **46**, 89–104, 1977.
45. M. J. Field, P. A. Bash, and M. Karplus, *A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations*, J. Comput. Chem., **11**, 700–733, 1990.
46. R. C. Walker, M. F. Crowley, and D. A. Case, *The implementation of a fast and accurate QM/MM potential method in Amber*, J. Comput. Chem., **29**, 1019–1031, 2008.
47. N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, and W. F. van Gunsteren, *Definition and testing of the GROMOS force-field versions 54A7 and 54B7*, Eur. Biophys. J., **40**, 843–856, 2011.
48. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules*, J. Am. Chem. Soc., **117**, 5179–5197, 1995.
49. A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguven, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, *All-Atom Empircal Potential for Molecular Modeling and Dynamics Studies of Proteins*, J. Phys. Chem., **102**, 3586–3626, 1998.

50. K. Meier, W. Thiel, and W. F. van Gunsteren, *On the effect of a variation of the force field, spatial boundary condition and size of the QM region in QM/MM MD simulations.*, J. Comput. Chem., **33**, 363–78, 2012.
51. C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren, *A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6*, J. Comput. Chem., **25**, 1656–1676, 2004.
52. W. E. Reiher, *Theoretical studies of hydrogen bonding*, PhD thesis, Harvard University, 1985.
53. R. E. Bulo, C. Michel, P. Fleurat-Lessard, and P. Sautet, *Multiscale Modeling of Chemistry in Water: Are We There Yet?*, J. Chem. Theory Comput., **9**, 5567–5577, 2013.
54. T. S. Hofer, B. R. Randolf, and B. M. Rode, *Influence of polarization and many body quantum effects on the solvation shell of Al(III) in dilute aqueous solution – extended ab initio QM/MM MD simulations*, Phys. Chem. Chem. Phys., **7**, 1382–1387, 2005.
55. S. Riahi, B. Roux, and C. N. Rowley, *QM/MM molecular dynamics simulations of the hydration of Mg (II) and Zn (II) ions*, Can. J. Chem., **91**, 1–7, 2013.
56. O. M. D. Lutz, C. B. Messner, T. S. Hofer, L. R. Canaval, G. K. Bonn, and C. W. Huck, *Computational Vibrational Spectroscopy of glycine in aqueous solution – Fundamental considerations towards feasible methodologies*, Chem. Phys., **435**, 21–28, 2014.
57. V. López-Canut, S. Martí, J. Bertrán, V. Moliner, and I. Tuñón, *Theoretical modeling of the reaction mechanism of phosphate monoester hydrolysis in alkaline phosphatase*, J. Phys. Chem. B, **113**, 7816–7824, 2009.
58. R. B. Murphy, D. M. Philipp, and R. A. Friesner, *A Mixed Quantum Mechanics (QM/MM) Method for Large-Scale Modeling of Chemistry in Protein Environments*, Chem. Phys. Lett., **21**, 1442–1457, 2000.
59. U. Pentikäinen, K. E. Shaw, K. Senthilkumar, C. J. Woods, and A. J. Mulholland, *Lennard-Jones Parameters for B3LYP/CHARMM27 QM/MM Modeling of Nucleic Acid Bases*, J. Chem. Theory Comput., **5**, 396–410, 2009.
60. D. Riccardi, G. Li, and Q. Cui, *Importance of van der Waals Interactions in QM/MM Simulations.*, J. Phys. Chem. B, **108**, 6467–78, 2004.
61. P. A. Bash, L. L. Ho, A. D. MacKerell Jr., D. Levine, and P. Hallstrom, *Progress toward chemical accuracy in the computer simulation of condensed phase reactions*, Proc. Natl. Acad. Sci. USA, **93**, 3698–3703, 1996.
62. M. Freindorf and J. Gao, *Optimization of the Lennard-Jones Parameters for a Combined Ab Initio Quantum Mechanical and Molecular Mechanical Potential Using the 3-21G Basis Set*, J. Comput. Chem., **17**, 386–395, 1996.
63. F. J. Luque, N. Reuter, A. Cartier, and M. F. Ruiz-López, *Calibration of the quantum/classical Hamiltonian in semiempirical QM/MM AM1 and PM3 methods*, J. Phys. Chem. A, **104**, 10923–10931, 2000.
64. M. Freindorf, Y. Shao, T. R. Furlani, and J. Kong, *Lennard-Jones parameters for the combined QM/MM method using the B3LYP/6-31G*/AMBER potential.*, J. Comput. Chem., **26**, 1270–8, 2005.
65. M. Dupuis, M. Aida, Y. Kawashima, and K. Hirao, *A polarizable mixed Hamiltonian model of electronic structure for micro-solvated excited states. I. Energy and gradients formulation and application to formaldehyde ($^1A_2$)*, J. Chem. Phys., **117**, 1242–1255, 2002.

66. Y. Tu and A. Laaksonen, *On the effect of Lennard-Jones parameters on the quantum mechanical and molecular mechanical coupling in a hybrid molecular dynamics simulation of liquid water*, J. Chem. Phys., **111**, 7519, 1999.
67. P. L. Cummins and J. E. Gready, *Coupled semiempirical molecular orbital and molecular mechanics model (QM/MM) for organic molecules in aqueous solution*, J. Comput. Chem., **18**, 1496–1512, 1997.
68. P. L. Cummins and J. E. Gready, *Coupled semiempirical quantum mechanics and molecular mechanics (QM/MM) calculations on the aqueous solvation free energies of ionized molecules*, J. Comput. Chem., **20**, 1028–1038, 1999.
69. M. L. Sánchez, M. A. Aguilar, and F. J. Olivares del Valle, *Study of solvent effects by means of averaged solvent electrostatic potentials obtained from molecular dynamics data*, J. Comput. Chem., **18**, 313–322, 1997.
70. M. L. Sánchez, M. E. Martín, M. A. Aguilar, and F. J. Olivares del Valle, *Solvent effects by means of averaged solvent electrostatic potentials: Coupled method*, J. Comput. Chem., **21**, 705–715, 2000.
71. M. E. Martín, M. A. Aguilar, S. Chalmet, and M. F. Ruiz-López, *An iterative procedure to determine Lennard-Jones parameters for their use in quantum mechanics/molecular mechanics liquid state simulations*, Chem. Phys., **284**, 607–614, 2002.
72. Y. Koyano, N. Takenaka, Y. Nakagawa, and M. Nagaoka, *On the Importance of Lennard-Jones Parameter Calibration in QM/MM Framework: Reaction Path Tracing via Free Energy Gradient Method for Ammonia Ionization Process in Aqueous Solution*, Bulletin of the Chemical Society of Japan, **83**, 486–494, 2010.
73. S. Chalmet and M. F. Ruiz-López, *New approaches to the description of short-range repulsion interactions in hybrid quantum/classical systems*, Chem. Phys. Lett., **329**, 154–159, 2000.
74. Y. Jin, E. R. Johnson, X. Hu, W. Yang, and H. Hu, *Contributions of Pauli repulsions to the energetics and physical properties computed in QM/MM methods.*, J. Comput. Chem., **34**, 2380–2388, 2013.
75. P. L. Freddolino, C. B. Harrison, Y. Liu, and K. Schulten, *Challenges in protein-folding simulations*, Nature Physics, **6**, 751–758, 2010.
76. H. Yu and W. F. van Gunsteren, *Accounting for polarization in molecular simulation*, Comput. Phys. Commun., **172**, 69–85, 2005.
77. A. Warshel, M. Kato, and A. V. Pisliakov, *Polarizable force fields: history, test cases, and prospects*, J. Chem. Theory Comput., **3**, 2034–2045, 2007.
78. P. E. M. Lopes, B. Roux, and A. D. Mackerell Jr., *Molecular modeling and dynamics studies with explicit inclusion of electronic polarizability. Theory and applications.*, Theoret. Chem. Acc., **124**, 11–28, 2009.
79. P. Cieplak, F. Dupradeau, Y. Duan, and J. Wang, *Polarization effects in molecular mechanical force fields.*, J. Phys.: Condens. Matter, **21**, 333102, 2009.
80. G. Babitzki, R. Denschlag, and P. Tavan, *Polarization Effects Stabilize Bacteriorhodopsin's Chromophore Binding Pocket: A Molecular Dynamics Study*, J. Phys. Chem. B, **113**, 10483–10495, 2009.
81. H. Luecke, B. Schobert, H. Richter, J. Cartailler, and J. K. Lanyi, *Structure of Bacteriorhodopsin at 1.55 Å Resolution*, J. Mol. Biol., **291**, 899–911, 1999.
82. G. Babitzki, G. Mathias, and P. Tavan, *The Infrared Spectra of the Retinal Chromophore in Bacteriorhodopsin Calculated by a DFT/MM Approach*, J. Phys. Chem. B, **113**, 10496–10508, 2009.

83. B. Rieff, S. Bauer, G. Mathias, and P. Tavan, *DFT/MM Description of Flavin IR Spectra in BLUF Domains*, J. Phys. Chem. B, **115**, 11239–11253, 2011.
84. S. Patel and C. L. Brooks, *CHARMM fluctuating charge force field for proteins: I parameterization and application to bulk organic liquid simulations*, J. Comput. Chem., **25**, 1–15, 2004.
85. V. M. Anisimov, G. Lamoureux, I. V. Vorobyov, N. Huang, B. Roux, and A. D. MacKerell, *Determination of electrostatic parameters for a polarizable force field based on the classical Drude oscillator*, J. Chem. Theory Comput., **1**, 153–168, 2005.
86. B. T. Thole, *Molecular polarizabilities calculated with a modified dipole interaction*, Chem. Phys., **59**, 341–350, 1981.
87. Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, *Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins*, J. Chem. Theory Comput., **9**, 4046–4063, 2013.
88. M. Masia, *Ab initio based polarizable force field parametrization*, J. Chem. Phys., **128**, 184107, 2008.
89. P. Tröster, K. Lorenzen, M. Schwörer, and P. Tavan, *Polarizable water models from mixed computational and empirical optimization*, J. Phys. Chem. B, **117**, 9486–9500, 2013.
90. M. J. Field, *Hybrid quantum mechanical/molecular mechanical fluctuating charge models for condensed phase simulations*, Mol. Phys., **91**, 835–846, 1997.
91. R. A. Bryce, R. Buesnel, I. H. Hillier, and N. A. Burton, *A solvation model using a hybrid quantum mechanical/molecular mechanical potential with fluctuating solvent charges*, Chem. Phys. Lett., **279**, 367, 1997.
92. I. H. Hillier, *Chemical reactivity studied by hybrid QM/MM methods*, J. Mol. Struct.: THEOCHEM, **463**, no. 1-2, 45–52, 1999.
93. D. P. Geerke, S. Thiel, W. Thiel, and W. F. van Gunsteren, *Combined QM/MM molecular dynamics study on a condensed-phase SN2 reaction at nitrogen: the effect of explicitly including solvent polarization*, J. Chem. Theory Comput., **3**, 1499–1509, 2007.
94. S. Caprasecca, S. Jurinovich, L. Viani, C. Curutchet, and B. Mennucci, *Geometry Optimization in Polarizable QM/MM Models: The Induced Dipole Formulation*, J. Chem. Theory Comput., **10**, 1588–1598, 2014.
95. M. A. Thompson and G. K. Schenter, *Excited States of the Bacteriochlorophyll b Dimer of Rhodopseudomonas viridis: A QM/MM Study of the Photosynthetic Reaction Center That Includes MM Polarization*, J. Phys. Chem., **99**, 6374–6386, 1995.
96. M. A. Thompson, *QM/MMpol: A Consistent Model for Solute/Solvent Polarization. Application to the Aqueous Solvation and Spectroscopy of Formaldehyde, Acetaldehyde, and Acetone*, J. Phys. Chem., **100**, 14492–14507, 1996.
97. Z. Lu and Y. Zhang, *Interfacing ab Initio Quantum Mechanical Method with Classical Drude Osillator Polarizable Model for Molecular Dynamics Simulation of Chemical Reactions*, J. Chem. Theory Comput., **4**, 1237–1248, 2008.
98. U. C. Singh and P. A. Kollman, *A Combined* Ab Initio *Quantum Mechanical and Molecular Mechanical Method for Carrying out Simulations on Complex Molecular Systems: Applications to the $CH_3Cl + CL^-$ Exchange Reaction and Gas Phase Protonation of Polyethers*, J. Comput. Chem., **7**, 718–730, 1986.

99. R. A. Bryce, M. A. Vincent, N. O. J. Malcolm, I. H. Hillier, and N. A. Burton, *Co-operative effects in the structuring of fluoride water clusters: Ab initio hybrid quantum mechanical/molecular mechanical model incorporating polarizable fluctuating charge solvent*, J. Chem. Phys., **109**, 3077, 1998.
100. H. Li and M. S. Gordon, *Polarization energy gradients in combined quantum mechanics, effective fragment potential, and polarizable continuum model calculations.*, J. Chem. Phys., **126**, 124112, 2007.
101. P. K. Biswas and V. Gogonea, *A polarizable force-field model for quantum-mechanical-molecular-mechanical Hamiltonian using expansion of point charges into orbitals*, J. Chem. Phys., **129**, 154108, 2005.
102. J. Gao, *Energy components of aqueous solution: insight from hybrid QM/MM simulations using a polarizable solvent model*, J. Comput. Chem., **18**, 1061–1071, 1997.
103. C. J. R. Illingworth, S. R. Gooding, P. J. Winn, G. A. Jones, G. G. Ferenczy, and C. A. Reynolds, *Classical polarization in hybrid QM/MM methods.*, J. Phys. Chem. A, **110**, 6487–97, 2006.
104. H. Houjou, Y. Inoue, and M. Sakurai, *Study of the Opsin Shift of Bacteriorhodopsin: Insight from QM/MM Calculations with Electronic Polarization Effects of the Protein Environment*, J. Phys. Chem. B, **105**, 867–879, 2001.
105. M. Wanko, M. Hoffmann, J. Frähmcke, T. Frauenheim, and M. Elstner, *Effect of polarization on the opsin shift in rhodopsins. 2. Empirical polarization models for proteins*, J. Phys. Chem. B, **112**, 11468–11478, 2008.
106. J. M. Olsen, K. Aidas, K. V. Mikkelsen, and J. Kongsted, *Solvatochromic shifts in uracil: a combined MD-QM/MM study*, J. Chem. Theory Comput., **6**, 249–256, 2009.
107. Y. Mo and J. Gao, *Polarization and charge-transfer effects in aqueous solution via ab initio QM/MM simulations.*, J. Phys. Chem. B, **110**, 2976–80, 2006.
108. Q. Wang and R. A. Bryce, *Accounting for non-optimal interactions in molecular recognition: a study of ion-π complexes using a QM/MM model with a dipole-polarisable MM region.*, Phys. Chem. Chem. Phys., **13**, 19401–8, 2011.
109. M. P. Gaigeot, R. Vuilleumier, M. Sprik, and D. Borgis, *Infrared spectroscopy of N-methylacetamide revisited by ab initio molecular dynamics simulations*, J. Chem. Theory Comput., **1**, 772–789, 2005.
110. H. W. Hugosson, A. Laio, P. Maurer, and U. Rothlisberger, *A comparative theoretical study of dipeptide solvation in water.*, J. Comput. Chem., **27**, 672–84, 2006.
111. J. VandeVondele, P. Tröster, P. Tavan, and G. Mathias, *Vibrational spectra of phosphate ions in aqueous solution probed by first-principles molecular dynamics.*, J. Phys. Chem. A, **116**, 2466–74, 2012.
112. G. Tabacchi, C. J. Mundy, J. Hutter, and M. Parrinello, *Classical polarizable force fields parametrized from ab initio calculations*, J. Chem. Phys., **117**, 1416–1433, 2002.

Coarse-Grain Water and Solutions: Hybrid AA/CG with Martini

Lars Schäfer

Lehrstuhl für Theoretische Chemie
Ruhr-University Bochum, 44780 Bochum, Germany

E-mail: lars.schaefer@rub.de

1 Introduction

One of the major bottlenecks of conventional all-atom (AA) molecular dynamics (MD) simulations of biomolecular systems comprised of hundreds of thousands of atoms is the very large computational effort involved. Even with efficient force fields for describing the pairwise interactions between the atoms, processes that can be directly simulated are limited to time- and length-scales up to a few μs and 20 - 30 nm, respectively. Of course, these boundaries are constantly being expanded, both due to hardware advancements (including GPUs and the design of dedicated special-purpose machines[1]) and the complementary development of MD software to actually exploit these in an optimal manner. Nevertheless, despite tremendous progress in the field, it is clear that brute-force all-atom MD alone will not suffice.

To overcome these limitations, efficient coarse-grained (CG) models have been developed.[2] In particle-based CG models, several atoms are grouped together into CG beads, thereby increasing computational efficiency by up to several orders of magnitude. Coarse-graining the solvent is particularly appealing, especially for biomolecular systems, because the majority of the computational effort is typically spend on computing interactions involving solvent molecules. CG approaches have been successfully applied to a wide range of biomolecular systems and processes. However, due to the unavoidable approximations that are inherent to coarse-graining, CG models are, in general, less accurate than their AA counterparts. This is the major limitation for their range of application. For example, for most CG models that are computationally highly efficient, it is very challenging to correctly describe conformational transitions and internal dynamics of complex biological macromolecules, including the making and breaking of secondary structure elements and folding/unfolding.

The aim of multiscale modeling is to trade off accuracy against efficiency by combining several levels of resolution. In hybrid methods, these levels of resolution (e.g., AA and CG) are present simultaneously in the simulation system, thus requiring direct interactions (coupling) between them. A hybrid simulation can for example be set-up by partitioning the system into different spatial regimes. In many cases, the processes of interest take place in a rather small part of the overall simulation system, such as ligand-receptor binding or the conformational change of a part of a biomolecule. In these cases, hybrid dual-scale AA/CG methods, in which an AA force field is used for the small subsystem of interest and coupled to a CG description of the surrounding, including the vast majority of the solvent molecules, can be expected to provide a substantial computational speed-up.

However, to judge the suitability of such approaches, it is critical to evaluate their accuracy by systematic comparison against fully atomistic simulations and experiment.

The aim of these lecture notes is to provide an overview over our work on developing dual-scale AA/CG hybrid methods to couple existing AA force fields with the widely-used Martini CG force field. The focus is on methodological aspects and their practical implementation in the GROMACS[3] MD code. Benefits and limitations of the current AA/CG approach – as well as possible ways to overcome the latter in the future – are discussed.

2 Brief Introduction to the CG-Martini Force Field

The Martini CG force field is one of the most widely-used CG models for biomolecular simulations.[4] Force field parameters, sample input files, scripts for setting up simulations, and tutorials can be found on the website `http://www.cgmartini.nl`. Initially developed for lipids,[5] Martini was later extended to other biomolecules such as proteins[6,7] and carbohydrates,[8] thus enabling a broad range of applications. The basic idea of the CG-Martini model is to employ a 4:1 mapping of AA to CG particles.[a] This means that on average, 4 atomistic heavy atoms (and their associated hydrogens) are effectively represented by 1 CG bead. The interactions between the CG beads are described by Lennard-Jones (6,12) potentials. Explicit Coulomb interactions are only present between fully charged beads, such as e.g. ionized amino acid side chains or charged (or zwitterionic) lipid headgroups, and are uniformly screened with a relative dielectric constant of $\epsilon_{\mathrm{r}}(\mathrm{CG/CG}) = 15$ to account for the lack of explicit dielectric polarizability of the CG model. The differences in polarity between the (uncharged) CG beads is reflected in the depth of their pairwise Lennard-Jones interaction potentials. These are not derived from standard combination rules, but were parameterized against experimental thermodynamic data, such as oil/water partitioning coefficients. The bonded parameters (bonds, angles, dihedrals) were largely obtained from inverting the respective populations obtained from atomistic simulations.

Several CG water models were developed to be used together with Martini. All these models employ a 4:1 mapping of atomistic water molecules to a single supra-molecular CG water, thus providing a substantial computational speed-up. The original water model[5] is a pure Lennard-Jones fluid, i.e., a single (very polar) Lennard-Jones bead with a mass of 72 au. It was parameterized such as to match the density of water at room temperature and the partitioning free energies between water and apolar medium of various model CG compounds that represent chemical building blocks of biomolecules (see above). One of the drawbacks of this original water model is that, due to the lack of partial charges, it does not have explicit dielectric properties and hence cannot screen the Coulomb interactions between charged particles. The polarizable Martini water (PW) model[9] contains partial charges that can readjust their positions in response to the electric field exerted by other (partial) charges in their environment. The PW model consists of three particles (Fig. 1, left): an uncharged central particle (W) for the Lennard-Jones interaction, and two particles with opposite partial charges (WP and WM with q = +0.46 and -0.46, respectively) that are connected to the central particle by a constraint of length 0.14 nm. These satellite particles, which do not interact through Lennard-Jones potentials, are within the excluded volume of the central particle. The WM-W-WP angle is modeled with a harmonic potential with

[a] This effective 4:1 mapping is applied on average and is by no means a strict rule – there are numerous exceptions.

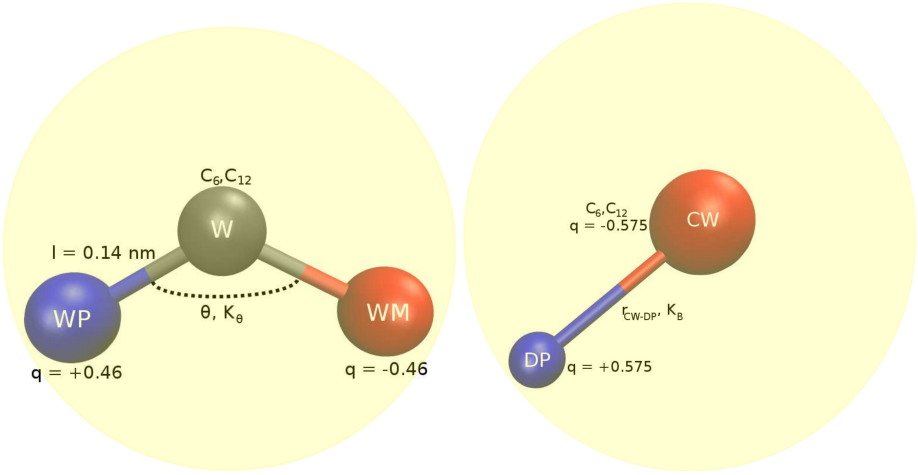


Figure 1: Polarizable coarse-grained water models. (Left) Polarizable Martini (PW) water model. (Right) Gromos big dipole water (BDW) water model.

equilibrium angle $\theta = 0$ and a weak force constant of $\mathrm{K}_\theta = 4.2\,\mathrm{kJ\,mol^{-1}\,rad^{-2}}$ and is thus flexible. Hence, the PW model accounts for orientational polarization. The Lennard-Jones interaction of the W bead, the partial charges, and K_θ were adjusted such as to reproduce as closely as possible the density and dielectric permittivity of water at room temperature. At 300 K, the PW model yields $\rho = 1043\,\mathrm{g/l}$ and $\epsilon = 76$. Thanks to the presence of explicit partial charges, ϵ_r could be reduced from 15 in standard Martini to 2.5 in PW.

The big multipole water (BMW) model[10,11] is another polarizable CG water model that can be used together with Martini. Just like PW, BMW is a 3-site CG water model that effectively represents 4 water molecules. But distinct from PW, it has a charged central W particle (q = -2), to which the two positively charged satellites (+1) are bonded. BMW is a rigid model with bonds (l = 0.12 nm) and angle ($\theta = 120°$) fixed and can thus be called a "super-SPC". In addition to a dipole, it also has a quadrupole moment. The Coulomb interactions are screened with $\epsilon_\mathrm{r} = 1.3$, and the van-der-Waals interactions are described with a soft Born-Mayer-Huggins potential instead of a steep Lennard-Jones 6,12 potential.

The third CG water model discussed in this lecture is the Gromos CG water model of Riniker et al.[12], also referred to as the big dipole water (BDW) model. A single BDW molecule represents 5 atomistic water molecules (instead of 4, as PW and BDW). As shown in Fig. 1, BDW consists of two particles, a central particle (CW) with a mass of 60 au that carries Lennard-Jones 6,12 parameters, and a dipole particle (DP) with a mass of 30 au. These two particles are oppositely charged (q = -/+ 0.575) and bonded to each other through a flexible half-attractive flat-bottom quartic potential, $\mathrm{V_{bond}(r_{CW-DP}) = 1/2\,K_B(r_{CW-DP} - r_0)^4}$, which acts beyond an inter-particle distance of $\mathrm{r_0} > 0.2\,\mathrm{nm}$ with a force constant of $\mathrm{K_B} = 2 \cdot 10^6\,\mathrm{kJ\,mol^{-1}\,nm^{-4}}$. Since the DP particle can oscillate around CW in response to the local field in the environment, the model is polarizable. The quartic bond makes the polarizability nonlinear. As for the other polarizable CG water models, the non-bonded interactions between particles within a single bead

are mutually excluded. Similar to PW, the BDW model uses $\epsilon_{\mathrm{r}} = 2.5$. Non-bonded interactions between BDW particles are taken into account up to a cut-off of 2.0 nm (instead of 1.2 nm for PW), which makes this model computationally more expensive. Although the BDW model was not parameterized together with Martini (i.e., there are no parameters for Lennard-Jones interactions between BDW and Martini beads), its use as a solvent in hybrid AA/CG modeling, as pioneered by Riniker and coworkers,[13,14] is discussed in this lecture.

The reasons for the computational efficiency of the CG-Martini model are threefold. First, the number of particles is significantly smaller (roughly by a factor of 10) than at the fully atomistic level, thus greatly reducing the number of pairwise interactions and forces that need to be computed. Second, direct long-range interactions are neglected, since all non-bonded interactions (and forces) are smoothly shifted to zero at a cut-off of 1.2 nm. Third, due to the lack of fast degrees of freedom such as hydrogen vibrations, the equations of motion can be integrated with larger time steps (of the order of 20 fs). Finally, depending on the particular system, lack of atomistic friction leads to a smoothed CG energy landscape on which sampling can be (but not necessarily always is) faster. Taken together, these effects provide a speed-up of up to a factor 1000 over atomistic models.

Finally, it should be mentioned that other polarizable CG water models have been devised, see e.g. Ref. 15 for a review. However, these were not developed to be used together with the Martini force field and are thus not covered in this lecture. In addition, very recently, an implicit solvent (dry) Martini force field was published.[16]

3 Direct Embedding of AA Solutes in CG Environment

Arguably, in terms of computational efficiency, completely replacing all atomistic solvent molecules by CG ones is very tempting, because it promises to provide the largest speed-up. However, at the same time, this rather drastic approach is also highly challenging due to the mismatch between the different levels of resolution, both in terms of structure (size of the CG beads *vs.* individual atoms) and thermodynamics (free energy difference between AA and CG force fields). In the AA/CG partitioning scheme presented in the following, the feasibility of such an approach was investigated. An AA solute of interest (e.g., (part of) a protein) is solvated in a purely CG environment. In addition, the molecules are assigned to be in the AA or CG subsystem *a priori*, and molecules can not change this pre-assigned resolution during the simulation.[b] Such a fixed-resolution set-up is conceptually similar to the conventional QM/MM methods and, similar to QM/MM, a major challenge lies in the coupling between the different level of resolution, that is, how to describe the cross-interactions between AA and CG. The original AA/CG-Martini coupling scheme was proposed by Rzepiela and coworkers.[17] The basic philosophy of their approach was to use the pure, unaltered force fields for the AA/AA and CG/CG interactions, respectively. The AA/CG cross-interactions were described at the level of the CG force field. To that aim, virtual CG interaction sites were introduced on groups of atomistic solute particles, according to the AA-to-CG mapping. Interactions between CG virtual sites, between CG virtual sites and AA particles, and direct interactions between AA and CG particles were

[b] As opposed to adaptive resolution methods, see lecture notes by Kurt Kremer in this issue.

excluded. The forces on the virtual CG sites due to the interactions with the CG surrounding were redistributed over the constituent AA particles such that the total momentum and torque are conserved. The advantage of this approach is that it is modular and straightforward, since it does not require any force field reparameterization. However, a drawback of this scheme, and also of related schemes such as the one suggested by Han and coworkers,[18] is the missing explicit dielectric screening of standard Martini CG water as a pure Lennard-Jones fluid. To screen the (otherwise too strong) Coulomb interactions between atomistic particles, Rzepiela et al. introduced a uniform relative dielectric permittivity, $\epsilon_{\mathrm{r}}(\mathrm{AA/AA}) > 1$. Different values of $\epsilon_{\mathrm{r}}(\mathrm{AA/AA})$ were found to be required for dialanine in water and in apolar solvent, showing that such a primitive uniform dielectric model is not suitable for describing large, heterogeneous systems such as proteins, in which the Coulomb interactions between protein atoms are not only screened by the solvent, but also by other protein atoms, depending on their location in the molecule.

Polarizable CG solvent models, such as the ones described above, enable an explicit electrostatic coupling between AA and CG subsystems. We explored whether explicit dielectric screening of AA/AA Coulomb interactions due to direct interactions with charged CG particles leads to an improved description of the electrostatics. Two AA/CG models are discussed in this lecture. First, our own hybrid model,[19] which uses the polarizable Martini (PW) water, and second, the model suggested by Riniker and coworkers,[13,14] which uses the BDW model. In both cases, different parameter sets of the Gromos force field were used for the AA part.[c] However, the AA/CG scheme is completely modular and can be readily extended to other atomistic force fields as well.

4 Details of the AA/CG Simulations

We will first discuss the details of the AA/CG model employing the PW and BMW water models, and later describe the (minor) differences that arise when the BDW model is used instead. All simulations were carried out with the GROMACS MD package (version 4.5),[3] in the NpT ensemble (p = 1 bar, T = 298 K) with an integration time step of 2 fs.[d] To put direct AA/CG electrostatic interactions into action (see Fig. 2), reaction field potentials $\mathrm{V_{rf}}$ and forces $\mathrm{F_{rf}}$ were used,

$$V_{rf} = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r}\left(r_{ij}^{-1} + k_{rf} r_{ij}^2 - c_{rf}\right) \tag{1}$$

and

$$F_{rf} = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r}\left(r_{ij}^{-2} - 2k_{rf} r_{ij}\right) \tag{2}$$

with

$$k_{rf} = r_c^{-3}(\epsilon_{rf} - \epsilon_r)(2\epsilon_{rf} + \epsilon_r)^{-1} \quad \text{and} \quad c_{rf} = r_c^{-1} + k_{rf} r_c^2.$$

[c] Strictly speaking, the Gromos force fields are united atom force fields, because the aliphatic H atoms are incorporated into the carbons. For simplicity, we refer to them as AA nevertheless.

[d] The CG degrees of freedom could be integrated with larger time steps (see Sec. 2). But no multiple time step integrator was used here.

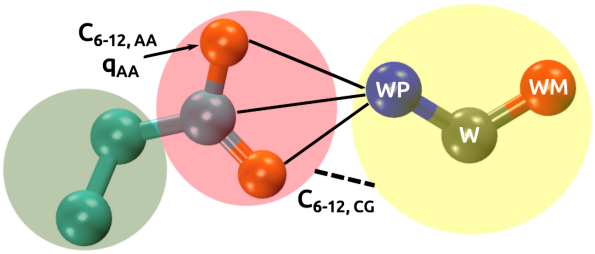


Figure 2: Basic concept of AA/CG coupling, illustrated for the interaction between an atomistic glutamate side chain (left) and a polarizable Martini CG water molecule (right). AA/CG Lennard-Jones interactions (dashed line) are described at the CG level by using CG virtual sites on the atomistic subsystem (green and red spheres represent the different CG-Martini bead types used for the apolar hydrocarbon chain and the carboxylic acid group, respectively). Partially charged particles in the AA and CG subsystems interact directly through Coulomb potentials (solid lines). The strength of this electrostatic coupling can be adjusted *via* $\epsilon_{\mathrm{r}}(\mathrm{AA/CG})$. The interactions within the AA and CG subsystems are described by the respective pure force fields, without alterations.

Here, $\mathrm{q_i}$ and $\mathrm{q_j}$ are the charges of the particles i and j, respectively; $\mathrm{r_{ij}}$ is the distance between the particles and $\mathrm{r_c} = 1.4\,\mathrm{nm}$ the cut-off radius. $\epsilon_0, \epsilon_{\mathrm{r}}$ and $\epsilon_{\mathrm{rf}} = 78$ are the vacuum, relative and reaction-field dielectric permittivities, respectively. For the AA electrostatic interactions, $\epsilon_{\mathrm{r}}(\mathrm{AA/AA}) = 1$, whereas the CG models have an $\epsilon_{\mathrm{r}}(\mathrm{CG/CG}) > 1$ (see Sec. 2). The strength of the AA/CG electrostatic coupling was systematically varied by changing $\epsilon_{\mathrm{r}}(\mathrm{AA/CG})$. The reaction-field electrostatics approach (instead of, e.g., particle-mesh Ewald (PME) long-range electrostatics) was used. This was done mainly for simplicity, because it enables to use different values of ϵ_{r} for the different classes of interactions – AA/AA, CG/CG, and AA/CG – simply by storing the respective $\mathrm{V_{rf}}$ and $\mathrm{F_{rf}}$ in look-up tables.[e] The values used for ϵ_{r} are 1 for the AA/AA interactions, and 2.5, 1.3, and 15 for the CG/CG interactions in PW, BMW, and standard (Lennard-Jones) Martini water, respectively. The different settings for the non-bonded interactions are summarized in Table 1. Charges on CG virtual sites were put to zero. In other words, CG virtual sites were only used to describe the AA/CG Lennard-Jones interactions (*via* the Martini CG potentials), but not for the Coulomb interactions, because the latter are accounted for by the direct AA/CG interactions.

The outlined coupling scheme potentially includes some degree of double-counting. For illustration, let's consider a polar amino acid side chain, e.g. Ser, in water. At the CG level, the Ser side chain (methanol) is modeled as an uncharged CG bead. Its polarity is reflected in the strongly attractive Lennard-Jones interaction with CG water. This Lennard-Jones interaction thus implicitly includes, to some degree, electrostatic interaction between the side chain OH-group and water. In the AA/CG hybrid scheme, explicit electrostatic interactions are added to these CG Lennard-Jones interactions (described at the CG level through virtual sites, see above), thus leading to a potential overestimation of the attractive interaction between the side chain and water. To investigate this effect, free energies of solvation, $\Delta \mathrm{G_{hyd}}$, of AA amino acid side chains in CG water have been calculated.[19] The

[e] The Lennard-Jones interactions were also tabulated.

		$\epsilon_{\mathbf{r}}$	Coulomb cut-off* [nm]	van der Waals cut-off [nm]
AA		1.0	1.4	1.4
AA/CG	std. W	n.a.	n.a.	1.2#
	PW	1.45&	1.4	1.2#
	BMW	1.1&	1.4	1.4
	BDW	2.3	2.0	2.0
CG/CG	std. W	n.a.	n.a.	1.2#
	PW	2.5	1.2	1.2#
	BMW	1.3	1.4	1.4+
	BDW	2.5	2.0	2.0

*reaction-field electrostatics, ϵ_{rf} = 78. #Shifted 0.9-1.2 nm.
&Adjusted parameter. +Born-Mayer-Huggins potential.

Table 1: Non-bonded interactions in AA/CG simulations.

key difference between the AA/CG scheme just described and the one proposed by Riniker et al.[13,14] for the BDW model is that in the latter, no CG virtual sites are used. Instead, the AA/CG Lennard-Jones interactions are derived from standard combination rules, i.e., geometric averaging for the Gromos force fields.

A specialty of the AA/CG simulations with polarizable CG water models is that care has to be taken to avoid a polarization catastrophe. Since the charged satellite beads do not carry Lennard-Jones parameters, they can in principle overlap with other particles. In the purely CG simulations, this does not occur, because the satellites are well within the excluded volume of the central particle. However, atomistic particles can penetrate into this excluded volume and charges can overlap, leading to numerical instabilities. To avoid this scenario, a short-range r^{-12}-repulsion was introduced between the CG water satellite particles and atomistic particles, with $\mathrm{C}_{12} = 10^{-7}\,\mathrm{kJ\,mol^{-1}\,nm^{12}}$ for the PW and BMW hybrid AA/CG models, and $\mathrm{C}_{12} = 7.7848 \cdot 10^{-10}\,\mathrm{kJ\,mol^{-1}\,nm^{12}}$ for the BDW AA/CG model. To test the influence of this choice, for PW, the free energy profile between the lysine and glutamate side chains (see below) was re-calculated with an extremely short-ranged repulsive potential ($\mathrm{C}_{12} = 10^{-12}\,\mathrm{kJ\,mol^{-1}\,nm^{12}}$). This did not change the free energy profile within the statistical errors.

5 Validation 1: Dimerization Free Energy Profiles

For judging the accuracy of any simulation model, in addition to structural parameters (which will be discussed in the next section), a thorough analysis of the thermodynamics is of paramount importance. In the following, potentials of mean force (or free energy profiles) of dimerization of AA solutes in CG water are discussed. These potentials of mean force (PMFs) are expected to be very sensitive towards the CG solvent model and the AA/CG coupling. As solutes, both charged and uncharged amino acid side chain analogs of different polarity were chosen, because these can report on the different contributions of the electrostatic interactions (dominant for the charged and polar solutes) and van der Waals interactions (dominant for the apolar solutes).

The details of the system set-ups and PMF calculations are described elsewhere.[19] Briefly, the applied constraint force method involves a set of distance constraint simulations, from which the PMF can be obtained from

$$PMF = -\int_{r}^{R_{max}} [\langle f_c \rangle_r + 2\, k_B T\, r^{-1}]\, dr, \tag{3}$$

where $\langle \mathrm{f_c} \rangle_\mathrm{r}$ is the average force on a constraint between the centers of mass of two solute molecules separated by distance r. At each constrained distance, 10 ns of MD sampling was carried for the time average $\langle \mathrm{f_c} \rangle_\mathrm{r}$. The r^{-1} term is an entropic correction that accounts for the fact that the configuration space volume grows with distance, i.e., gives a $-\,\mathrm{k_B T}\,\ln(4\pi \mathrm{r}^2)$ contribution to the free energy. This is obtained by integrating over the centrifugal force $2\,\mathrm{k_B T}\,\mathrm{r}^{-1}$. The PMF at $\mathrm{R_{max}}$, the maximum distance used, is (arbitrarily) set to zero. To ensure that this distance cut-off is sufficient, sufficiently large simulation boxes were used, so that solute-solute distances up to $\mathrm{R_{max}} = 2.2\,\mathrm{nm}$ could be covered.

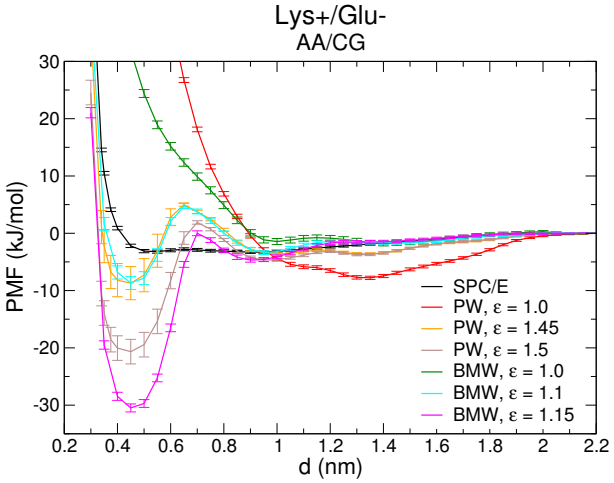


Figure 3: PMF of atomistic $\mathrm{Lys^+/Glu^-}$ ion pair solvated in polarizable CG water models PW and BMW. The strength of the AA/CG electrostatic interactions is varied with $\epsilon_\mathrm{r}(\mathrm{AA/CG})$.

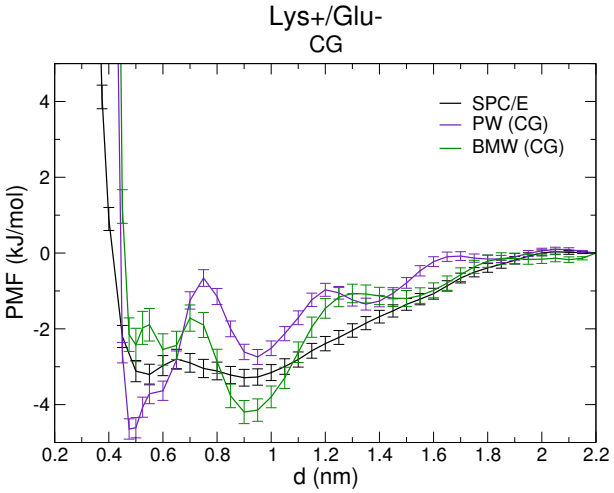


Figure 4: Reference PMFs of $\mathrm{Lys}^{+}/\mathrm{Glu}^{-}$ at the fully AA level (SPC/E water), and at the fully CG level.

The PMFs of the charged lysine/glutamate ($\mathrm{Lys}^{+}/\mathrm{Glu}^{-}$) ion pair are shown in Fig. 3 for the AA/CG simulations. For reference, Fig. 4 shows the corresponding PMFs obtained from fully atomistic or fully coarse-grained simulations of the same system. In Fig. 3 clearly shows the strikingly pronounced influence of the strength of the electrostatic AA/CG coupling on the free energy profiles, with strong changes in the PMF even within a rather small range of $\epsilon_{\mathrm{r}}(\mathrm{AA/CG})$. Weak AA/CG electrostatic coupling (large $\epsilon_{\mathrm{r}}(\mathrm{AA/CG})$) yields significantly too deep contact minima, whereas full coupling ($\epsilon_{\mathrm{r}}(\mathrm{AA/CG}) = 1.0$, red and green curves for PW and BMW, respectively) even leads to an unphysical repulsion between the oppositely charged ions. A compromise could be obtained with intermediate coupling strengths ($\epsilon_{\mathrm{r}}(\mathrm{AA/CG})$ = 1.45 and 1.1 for PW and BMW, respectively). However, these still have a too deep contact minimum and a significant barrier towards the solvent-separated minima.

For comparison, the PMFs of the fully AA or fully CG simulations display much shallower minima and smaller barriers, as shown in Fig. 4. As compared to the fully atomistic PMF in SPC/E water, the CG PMFs display a number of minima and maxima up to long distances. This over-structuring, which is also present in the AA/CG PMFs, is due to the granularity of the CG solvent (1 CG solvent bead represents 4 atomistic waters) and is slightly more pronounced for PW than for BMW, because the latter uses a softer Born-Mayer-Huggins potential instead of a (too steep) r^{-12} term to model the Pauli repulsion. A notable feature of the fully CG PMF obtained with the BMW model (green curve) is that the solvent-separated minimum is the global minimum, at odds with the PMFs in SPC/E and PW.

In addition to using the PW and BMW CG water models in the hybrid AA/CG simulations, the $\mathrm{Lys}^{+}/\mathrm{Glu}^{-}$ PMF was also calculated with the hybrid model suggested by Riniker and coworkers.[13,14] In their model, the BDW water model is used, with $\epsilon_{\mathrm{r}}(\mathrm{AA/CG})$ = 2.3. As described above, mixed AA/CG Lennard-Jones interactions are described using standard combination rules instead of CG virtual sites. Fig. 5 shows that

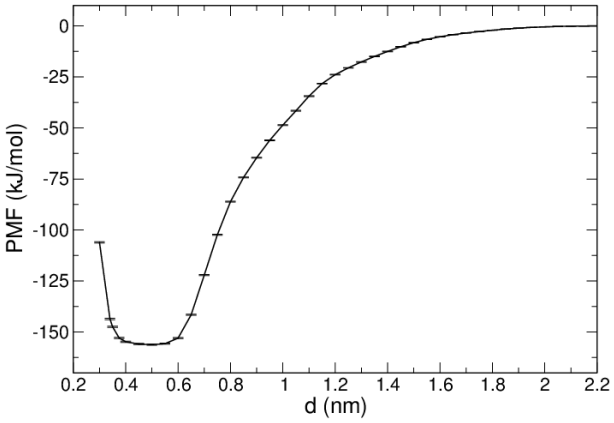


Figure 5: PMF of $\mathrm{Lys}^{+}/\mathrm{Glu}^{-}$ in BDW, using the hybrid AA/CG model proposed by Riniker et al.[13,14]

this model severely overestimates the depth of the minimum, by about 150 kJ/mol. Based on the above results, this can be attributed to the rather weak electrostatic coupling between AA and CG. The over-stabilization of salt bridges associated with this too deep minimum likely explains the previous observation[14] that in AA/CG simulations of atomistic proteins in BDW, the number of intra-protein hydrogen bonds was slightly larger than in the corresponding fully atomistic simulations. As further discussed below, this effect does not necessarily lead to completely unstable protein structures in an MD simulation. By contrast, since charged residues tend to be located at the protein surface, the formation of tight salt-bridges and hydrogen bonds between surface residues can lead to the formation of a cage-like network that can stabilize the protein structure in a conformation close to the starting structure. Thus, although the changes in terms of structure (deviation from the starting structure or change in number of hydrogen bonds/salt bridges) can be rather modest, possibly suggesting only minor deviations of the underlying energetics, only the PMF can provide the required quantitative picture.

The PMFs of the $\mathrm{Glu}^{-}/\mathrm{Glu}^{-}$ pair obtained from the hybrid AA/CG simulations and the reference PMFs from the fully AA or CG simulations are shown in Fig. 6 and 7, respectively. In these calculations, $\epsilon_{\mathrm{r}}(\mathrm{AA/CG})$ = 1.45 and 1.1 were chosen for PW and BMW, respectively, because these values yielded the most reasonable (albeit not perfect) results for $\mathrm{Lys}^{+}/\mathrm{Glu}^{-}$ (see Fig. 3). At the fully atomistic and fully CG level, repulsive PMFs are obtained (Fig. 7), as expected for an anion/anion pair. By contrast, the hybrid AA/CG PMFs (Fig. 6) have artificial minima at short distances. This can be remedied by increasing $\epsilon_{\mathrm{r}}(\mathrm{AA/CG})$ (magenta curve). However, as shown in Fig. 3, this weaker coupling leads to a strong over-stabilization of ± charge ion pairs. Similar effects are also expected for ions other than Lys^{+} and Glu^{-}. Although not systematically investigated here, these effects are probably worse for larger and more localized charges.

In summary, the PMFs of the studied ion pairs show that the free energy profiles obtained at the hybrid AA/CG level do strongly depend on the strength of the electrostatic interactions between the AA and CG subsystems, and can significantly differ from the reference PMFs. It appears to be very difficult to obtain good agreement with the AA

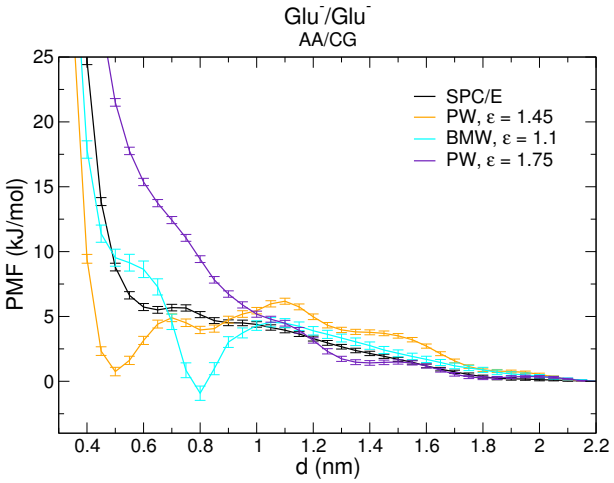


Figure 6: PMF of atomistic $\mathrm{Glu}^-/\mathrm{Glu}^-$ pair solvated in polarizable CG water models PW and BMW.

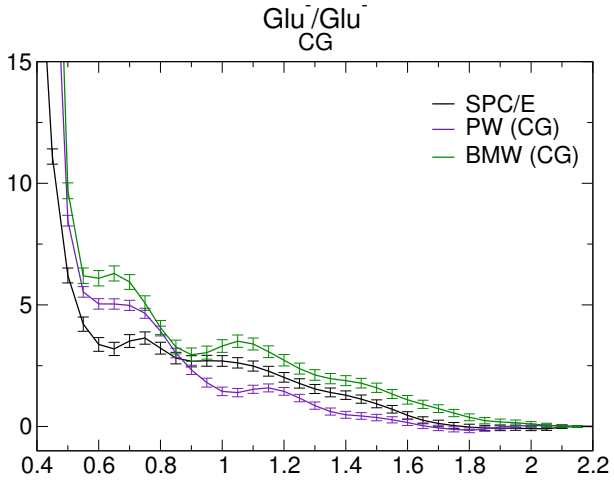


Figure 7: Reference PMFs of $\mathrm{Glu}^-/\mathrm{Glu}^-$ at the fully AA level (SPC/E water), and at the fully CG level.

reference PMFs by using $\epsilon_\mathrm{r}(\mathrm{AA/CG})$ as the only adjustable parameter. Generally, the free energy profile between two solute molecules does not only reflect the direct interactions between them, but is also influenced by multibody effects due to solute-solvent and solvent-solvent interactions. We attribute the shape of the hybrid AA/CG PMFs to the response of the polarizable CG water molecules to a particular solute configuration. For too strong interactions ($\epsilon_\mathrm{r}(\mathrm{AA/CG}) = 1$), this solvent over-polarization dominates the picture and can even lead to completely unreasonable PMFs, such as the effective repulsion between ± charge ion pairs (Fig. 3) or the attraction between two ions of like charge (Fig. 6).

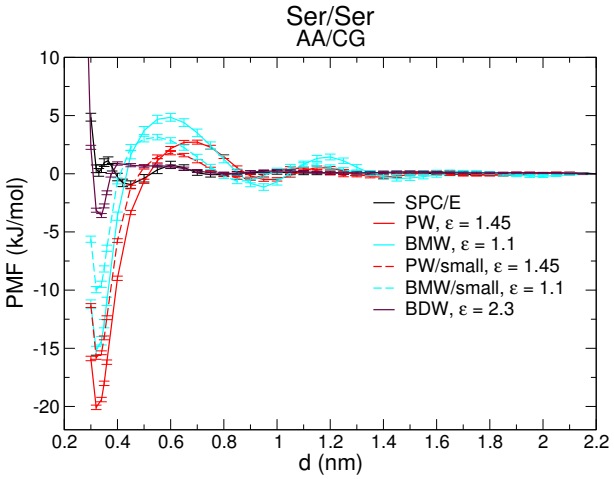


Figure 8: PMF of atomistic Ser/Ser (methanol/methanol) pair solvated in polarizable CG water models PW, BMW, and BDW. The dashed curves (PW/small and BMW/small) were obtained with a smaller CG bead ($\sigma_{\mathrm{LJ}} = 0.43\,\mathrm{nm}$ instead of 0.47 nm, and the well-depth ϵ_{LJ} was scaled by 0.75).

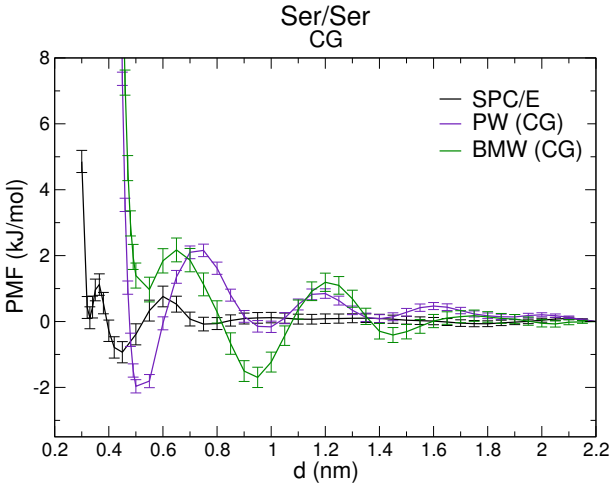


Figure 9: Reference PMFs of Ser/Ser at the fully AA level (SPC/E water), and at the fully CG level.

Next, we turn to solutes that do not carry a net charge. First, the Ser/Ser (methanol/methanol) pair is discussed as a representative example for a polar solute. Subsequently, Val/Val (propane/propane) was investigated as an example for an apolar solute. Note that in the Gromos (53a6) force field, Val is represented by 3 united-atom carbons that do not carry any partial charges. Hence, in this case, there are no solute-solvent electrostatic interactions at all. Figures 8 and 9 show the PMFs at the hybrid AA/CG and fully AA and fully CG levels, respectively.

The AA PMF (black curve) has two minima at short distances as a consequence of the peculiar hydrogen-bond arrangement of the methanol OH-groups. The hybrid AA/CG PMFs obtained with the PW and BMW CG water models display significantly too deep contact minima and too high desolvation barriers. The latter are partly due to the coarser granularity of the solvent, as discussed above. Notably, unlike for the ion pair (Fig. 5), the BDW CG model only slightly overestimates the depth of the minimum and, among the three CG water models, is closest to the atomistic reference. In addition, there is no pronounced desolvation barrier, despite its coarse granularity (1 BDW molecule represents 5 atomistic waters). These results speak in favor of using standard combination rules instead of CG virtual sites for the Lennard-Jones AA/CG coupling. None of the hybrid AA/CG PMFs captures the double-minimum at short distances, though. A peculiarity of the Ser side chain is its small size. Despite being comprised of only two heavy atoms (C-OH), it is, at the level of the CG-Martini force field, represented by a Lennard-Jones bead that has the same size as the ones as used for 4:1-mapped groups, i.e., $\sigma_{\mathrm{LJ}} = 0.47\,\mathrm{nm}$. The too large size of the Ser CG bead is clearly seen also in the fully CG PMFs (Fig. 9).[f] Since the Lennard-Jones AA/CG interactions are treated at the level of of the CG-Martini force field (*via* CG virtual sites), this leads to a too large excluded volume around the Ser side chain. To investigate this issue, the size of the CG Lennard-Jones bead was reduced to $\sigma_{\mathrm{LJ}} = 0.43\,\mathrm{nm}$ and the well-depth ϵ_{LJ} was scaled by 0.75. Indeed, the resulting PMFs (dashed lines in Fig. 8) have less deep minima and smaller barriers, although the stability of the contact pair is still overestimated.

Finally, Figures 10 and 11 show the PMFs of the Val/Val pair as an apolar solute. The contact minima in the AA/CG PMFs are slightly too deep, again with the exception of BDW that is in good agreement with SPC/E. The PMF from a hybrid AA/CG simulation with standard CG-Martini water (blue curve) is in this case similar to PW due to the absence of electrostatic AA/CG interactions. A similar agreement between full atomistic and hybrid AA/CG PMFs was also observed for the Phe/Phe (toluene/toluene) pair.[19]

In summary, the PMFs show that the investigated AA/CG coupling schemes do not provide a realistic description of the dimerization of charged solutes in water. For polar solutes, the situation is significantly improved, especially with BDW, although this CG water model most severely overestimates the stability of salt bridges (Fig. 5). Apolar solutes in water can be readily modeled at the AA/CG level.

6 Validation 2: Structure

In this section, we will demonstrate and discuss the pitfalls associated with solely focusing on structural parameters (instead of (free) energies) for evaluating the quality of a model. Fig. 12 shows the RMS deviation of ubiquitin from its starting structure in the course of a hybrid AA/CG MD simulation with the polarizable BDW water model, as compared to the same simulation carried out at the fully AA level in SPC water. Such simulations were previously carried out for other small proteins in water by Riniker et al.[14] A somewhat more detailed analysis is presented in Fig. 13, which shows the secondary structure plots.

[f] Again, the BMW model underestimates the free energy of the contact minimum with respect to the solvent-separated minimum, as also found for other solutes (Figures 4, 11).

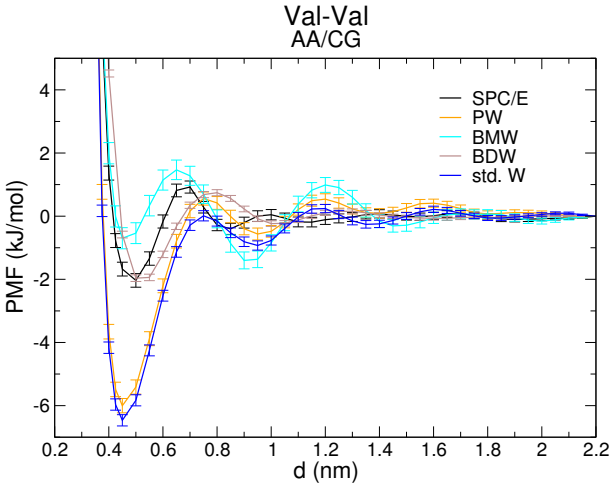


Figure 10: PMF of atomistic Val/Val pair solvated in polarizable CG water models PW, BMW, and BDW as well as in standard (Lennard-Jones) Martini water.

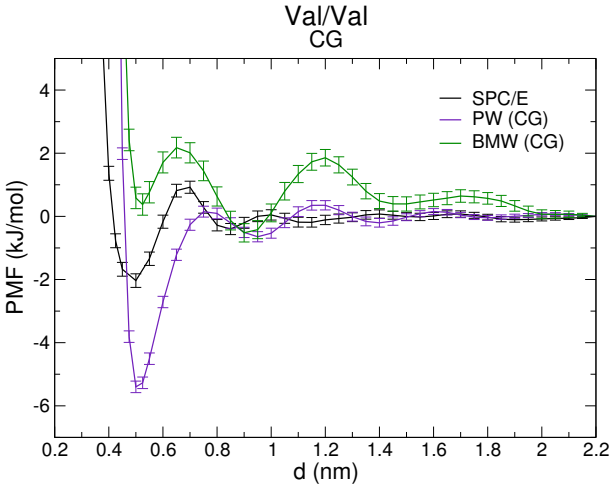


Figure 11: Reference PMFs of Val/Val at the fully AA level (SPC/E water), and at the fully CG level.

Surprisingly, despite the strong over-stabilization of salt bridges obtained with this AA/CG scheme (by more than 150 kJ/mol for the $\mathrm{Lys}^{+}/\mathrm{Glu}^{-}$ pair, Fig. 5), the RMSD is within reasonable bounds, and the secondary structure elements are stable on the simulation time scale. However, it is evident that the RMSD fluctuations are significantly reduced in the AA/CG simulation. Closer inspection of the simulation trajectory revealed that solvent-exposed charged and polar residues on the ubiquitin surface formed a network of salt bridges and hydrogen bonds that, once formed, did not break up again during the simulation. This over-stabilization locked the protein in a conformation close to the start-

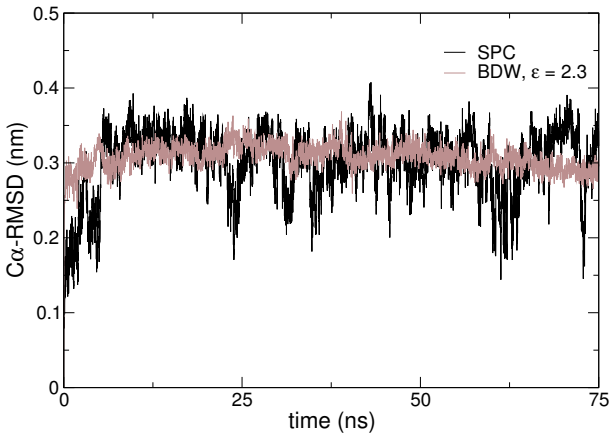


Figure 12: Root-mean square deviation (RMSD) of ubiquitin with respect to the starting structure (X-ray structure, PDB 1UBI).

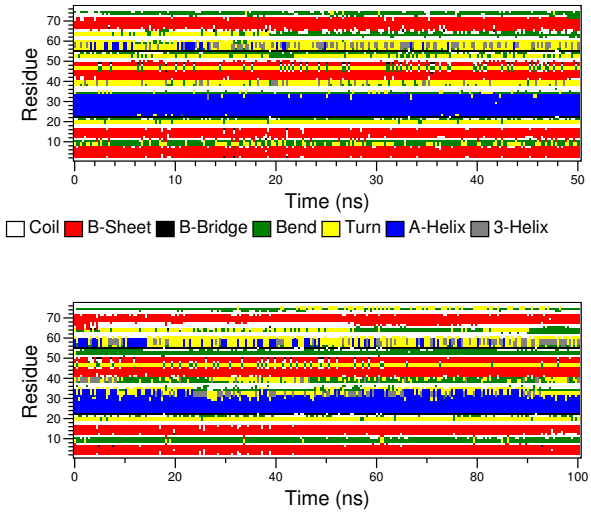


Figure 13: Secondary structure of ubiquitin during fully atomistic MD simulation in SPC (top) and during hybrid AA/CG simulation in BDW water with $\epsilon_r(\mathrm{AA/CG})$ = 2.3 (bottom).

ing structure – a kinetic trap from which there was no escape on the simulation time scale, because for structural rearrangements to occur, these interactions need to transiently break up.

As a second example to illustrate this structural over-stabilization, Fig. 14 shows the secondary structure plots of the GCN4-p1 peptide. At the fully atomistic level, with the

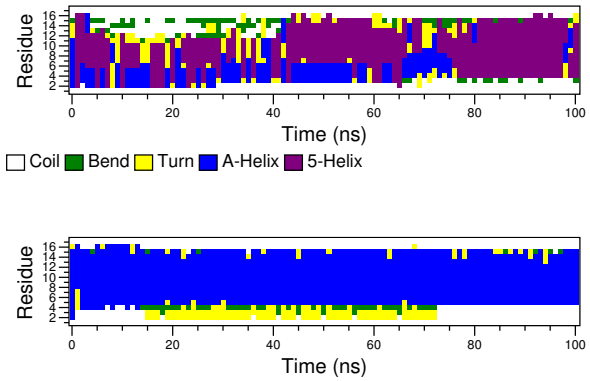


Figure 14: Secondary structure of GCN4-p1 peptide (PBD 2OVN) during fully atomistic MD simulation in SPC (top) and during hybrid AA/CG simulation in PW water with $\epsilon_{\mathrm{r}}(\mathrm{AA/CG})$ = 1.45 (bottom).

Gromos force field (43a2 parameter set in this case) and SPC water, the GCN4-p1 peptide dynamically changes between α- and π-helical conformations. At the hybrid AA/CG level, only α-helical conformations are sampled on the simulation time scale.[g]

Fig. 15 shows that this reduced flexibility is due to the formation of a very stable hydrogen-bond between Tyr17 and Glu22 (green curve) and a salt bridge between Glu22 and Arg25 (blue). Once formed, these interactions did not break up on the time scale of the AA/CG simulation. This was not the case in the fully atomistic reference simulation (black and red curves).

7 Summary and Conclusions

As discussed in this lecture, the applied hybrid AA/CG schemes in which an atomistic solute is directly solvated in CG water can be reasonably accurate for apolar solutes in water, and for solutes in an apolar medium. Polar and, even more so, charged solutes in water represent a major challenge, however. This can, in some sense, be attributed to the nature of CG water: As most CG water models, the PW, BDW, and BMW models have been designed to describe bulk water. Thus, *by construction*, these CG models can not be expected to accurately capture local solvation effects, which play a key role at interfaces or molecular surfaces. In particular, they can not form hydrogen bonds. In this sense, the CG water models could thus be considered to be inherently incompatible with AA models.[h] Possible remedies to this dilemma could be the introduction of explicit hydrogen-bond mimicking

[g] PW was used as CG water model here. Similar results were obtained with the BDW and BMW water models (not shown).

[h] Recently, an electrostatic coupling AA/CG method employing ELBA water, a more fine-grained CG water that describes a single water molecule as a permanent point dipole, has been successfully applied to study proteins.[20] Solute-solute PMFs have not been published for this model, though.

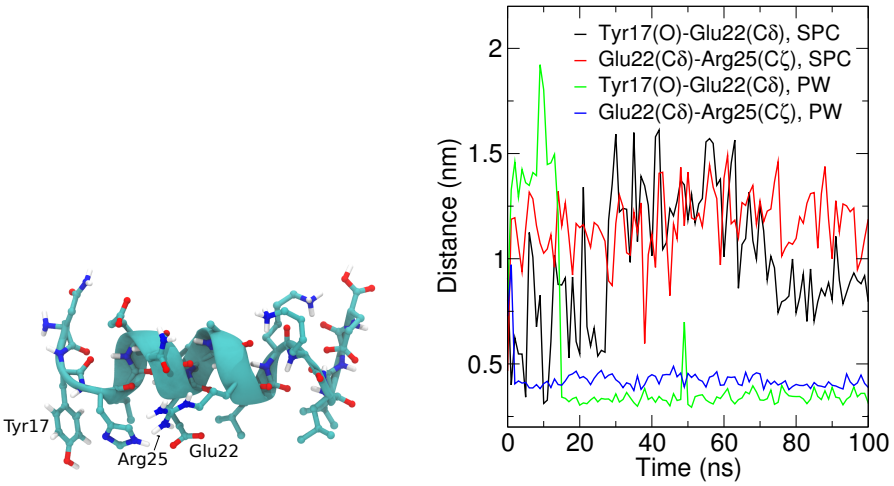


Figure 15: (Left) Structure of GCN4-p1 peptide. (Right) Distances between atoms in indicated residues.

AA/CG energy terms, which however is not particularly elegant and would require extensive (re-)parameterization. Another possibility could be to use different non-bonded interactions for different groups of atoms or residue pairs, depending on their charge/polarity. Although technically straightforward (through tabulated potentials), the use of different relative dielectric constants for the Coulomb interactions between different groups of atoms in the same system is difficult to justify. Finally, atomistic solvent shells could be introduced to facilitate coupling by shifting the AA/CG boundary further away from the solute of interest. Several researchers have suggested such schemes, both at the fixed-resolution level[21,22] and in adaptive resolution schemes.[23,24] However, when devising such atomic solvation layer methods, it should be kept in mind that computational speed-up is *the* main motivation for multiscale modeling. For the AA/CG simulations described in this lecture, the speed-up over the same simulations at the fully AA level was a factor of ca. 10. The use of multiple time step integrators would further boost this efficiency gain by another order of magnitude, since the CG degrees of freedom, which still represent the major part of the non-bonded interactions that determine the overall computational effort, could be integrated with 10 times larger time steps. The use of many atomic-level solvation layers[i] significantly diminishes the computational speed-up that can be achieved as compared to a fully atomistic simulation in a small, optimally-shaped periodic box. For the efficiency gain provided by a hybrid model to outplay the inherent loss in accuracy, it should enable to explore time and length scales that are otherwise not accessible to conventional atomistic simulations. This lecture highlighted some of the challenges that need to be overcome on the way toward hybrid AA/CG models that are highly computationally efficient, yet at the same time accurate enough to answer biological questions.

[i] Which in general might be required, because perturbation of water molecules due to the influence of an AA/CG boundary can be rather long-ranged, up to 1 nm.

Acknowledgments

The author thanks Alexander Kuhn for generating some of the Figures. This work was supported by the Deutsche Forschungsgemeinschaft (Cluster of Excellence RESOLV (EXC 1069), Emmy Noether grant to L.S.).

References

1. David E. Shaw et al., *Anton, a special-purpose machine for molecular dynamics simulation*, Commun. ACM, **51**, no. 7, 91, Jul 2008.
2. G. A Voth, (Ed.), *Coarse-Graining of Condensed Phase and Biomolecular Systems*, CRC Press / Taylor and Francis, Boca Raton, FL, 2009.
3. Sander Pronk, Szilárd Páll, Roland Schulz, Per Larsson, Pär Bjelkmar, Rossen Apostolov, Michael R. Shirts, Jeremy C. Smith, Peter M. Kasson, David van der Spoel, Berk Hess, and Erik Lindahl, *GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit*, Bioinformatics, **29**, no. 7, 845–854, 2013.
4. Siewert J. Marrink and D. Peter Tieleman, *Perspective on the Martini model*, Chem. Soc. Rev., **42**, no. 16, 6801–6822, Aug 2013.
5. Siewert J. Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H. de Vries, *The MARTINI force field: coarse grained model for biomolecular simulations*, J. Phys. Chem. B, **111**, no. 27, 7812–7824, Jul 2007.
6. Luca Monticelli, Senthil K. Kandasamy, Xavier Periole, Ronald G. Larson, D. Peter Tieleman, and Siewert J. Marrink, *The MARTINI coarse-grained force field: extension to proteins*, J. Chem. Theory Comput., **4**, no. 5, 819–834, 2008.
7. Djurre H. de Jong, Gurpreet Singh, W.F. Drew Bennett, Clement Arnarez, Tsjerk A. Wassenaar, Lars V. Schäfer, Xavier Periole, D. Peter Tieleman, and Siewert J. Marrink, *Improved parameters for the MARTINI coarse-grained protein force field*, J. Chem. Theory Comput., **9**, no. 1, 687–697, 2012.
8. Cesar A. López, Andrzej J. Rzepiela, Alex H. de Vries, Lubbert Dijkhuizen, Philippe H. Hünenberger, and Siewert J. Marrink, *MARTINI coarse-grained force field: extension to carbohydrates*, J. Chem. Theory Comput., **5**, no. 12, 3195–3210, 2009.
9. Semen O. Yesylevskyy, Lars V. Schäfer, Durba Sengupta, and Siewert J. Marrink, *Polarizable water model for the coarse-grained MARTINI force field*, PLoS Comput. Biol., **6**, no. 6, e1000810, June 2010.
10. Zhe Wu, Qiang Cui, and Arun Yethiraj, *A new coarse-grained model for water: the importance of electrostatic interactions*, J. Phys. Chem. B, **114**, no. 32, 10524–10529, 2010.
11. Zhe Wu, Qiang Cui, and Arun Yethiraj, *A New Coarse-Grained Force Field for Membrane–Peptide Simulations*, J. Chem. Theory Comput., **7**, no. 11, 3793–3802, 2011.
12. Sereina Riniker and Wilfred F. van Gunsteren, *A simple, efficient polarizable coarse-grained water model for molecular dynamics simulations*, J. Chem. Phys., **134**, no. 8, 084110, Feb 2011.
13. Sereina Riniker and Wilfred F. van Gunsteren, *Mixing coarse-grained and fine-grained water in molecular dynamics simulations of a single system*, J. Chem. Phys., **137**, no. 4, 44120, July 2012.

14. Sereina Riniker, Andreas P. Eichenberger, and Wilfred F. van Gunsteren, *Solvating atomic level fine-grained proteins in supra-molecular level coarse-grained water for molecular dynamics simulations*, Eur. Biophys. J., **41**, no. 8, 647–661, Aug. 2012.
15. Helgi I. Ingólfsson, Cesar A. López, Jaakko J. Uusitalo, Djurre H. de Jong, Srinivasa M. Gopal, Xavier Periole, and Siewert J. Marrink, *The power of coarse graining in biomolecular simulations*, WIREs Comput. Mol. Sci., **4**, no. 3, 225–248, 2014.
16. Clément Arnarez, Jaakko J. Uusitalo, Marcelo Fabricio Masman, Helgi I. Ingólfsson, Djurre H. de Jong, Manuel N. Melo, Xavier Periole, Alex H. de Vries, and Siewert J. Marrink, *Dry MARTINI, a coarse-grained force field for lipid membrane simulations with implicit solvent*, J. Chem. Theory Comput., **11**, no. 1, 260–275, 2014.
17. Andrzej J. Rzepiela, Martti Louhivuori, Christine Peter, and Siewert J. Marrink, *Hybrid simulations: combining atomistic and coarse-grained force fields using virtual sites*, Phys. Chem. Chem. Phys., **13**, no. 22, 10437–10448, Jun 2011.
18. Wei Han, Cheuk-Kin Wan, Fan Jiang, and Yun-Dong Wu, *PACE force field for protein simulations. 1. Full parameterization of version 1 and verification*, J. Chem. Theory Comput., **6**, no. 11, 3373–3389, 2010.
19. Tsjerk A. Wassenaar, Helgi I. Ingólfsson, Marten Prieß, Siewert J. Marrink, and Lars V. Schäfer, *Mixing MARTINI: electrostatic coupling in hybrid atomistic-coarse-grained biomolecular simulations*, J. Phys. Chem. B, **117**, no. 13, 3516–3530, Apr 2013.
20. Mario Orsi, Wei Ding, and Michail Palaiokostas, *Direct Mixing of Atomistic Solutes and Coarse-Grained Water*, J. Chem. Theory Comput., **10**, no. 10, 4684–4693, Oct 2014.
21. Sereina Riniker, Andreas P. Eichenberger, and Wilfred F. van Gunsteren, *Structural effects of an atomic-level layer of water molecules around proteins solvated in supra-molecular coarse-grained water*, J. Phys. Chem. B, **116**, no. 30, 8873–8879, Aug. 2012.
22. Leonardo Darré, Alex Tek, Marc Baaden, and Sergio Pantano, *Mixing atomistic and coarse grain solvation models for MD simulations: let WT4 handle the bulk*, J. Chem. Theory Comput., **8**, no. 10, 3880–3894, 2012.
23. Matej Praprotnik, Luigi Delle Site, and Kurt Kremer, *Multiscale simulation of soft matter: From scale bridging to adaptive resolution*, Annu. Rev. Phys. Chem., **59**, 545–571, 2008.
24. Steven O. Nielsen, Rosa E. Bulo, Preston B. Moore, and Bernd Ensing, *Recent progress in adaptive multiscale molecular dynamics simulations of soft matter*, Phys. Chem. Chem. Phys., **12**, no. 39, 12401–12414, 2010.

Adaptive Resolution Simulation

Raffaello Potestio[1], Aoife Fogarty[1], Christine Peter[2], and Kurt Kremer[1]

[1] Max-Planck-Institut für Polymerforschung, Ackermannweg 10, 55128 Mainz, Germany

[2] Department of Chemistry, Universität Konstanz, Universitätsstr. 10, 78464 Konstanz, Germany

E-mail: {*potestio, fogarty, kremer*}*@mpip-mainz.mpg.de, christine.peter@uni-konstanz.de*

While over the last few decades scale-bridging computer simulations have become a fundamental tool in the field of soft matter science, the combination of several levels of resolution within one simulation is based on more recent developments. For the employment of coarse-grained models, in which only a subset of the system's degrees of freedom is retained; for an effective and insightful use of these simplified models, though, an appropriate parametrization of the interactions is of crucial importance. However, in many cases the removal of fine-grained details in a specific, small region of the system would destroy relevant features; this circumstance can be approached with the use of dual-resolution simulation methods, where a subregion of the system is described with high resolution, and a coarse-grained representation is employed in the rest of the simulation domain. Two related computational methods implementing this approach are the AdResS and H-AdResS simulation schemes, which will be discussed here. This contribution is based on a recent review[1] and some new developments[2].

1 Introduction

Since the pioneering work carried out by Berni Alder[3] in the 1950ies, *in silico* experiments, such as Molecular Dynamics (MD) or Monte Carlo (MC) simulations, allowed researchers to obtain major advancements in the understanding of systems with many degrees of freedom. Particularly, during the last few decades the improved models, increasing accuracy of force-fields, the improvement of the algorithms, and the steady boost of computer power made it possible to perform insightful simulations of a broad variety of systems of increasing size and complexity, ranging from simple liquids -composed by idealized, point-like molecules interacting *via* simple potentials- to biomolecules. Nonetheless, the amount of available computational resources can be insufficient to simulate, for a physically meaningful time, even the simplest nontrivial macromolecule. Also the limited accuracy of atomistic force fields as well as the rather limited transferability of both atomistic and especially coarse-grained force fields is a permanent source of problems and concerns. It is often the case, in fact, that 'interesting' phenomena in these systems occur on very long time-scales: a simple example of this is provided by the diffusion of a polymer in a melt[4,5]; the same behavior can observed in conformational changes of proteins[6–11], at least in those cases in which the force field provides a good approximation to the real atomistic interactions.

At the same time, in many cases the massive amount of data that are produced in a simulation is composed in large part by non-useful information. A prototypical example is given by the solvent: the water molecules that solvate a protein or a membrane are typically discarded from the analysis that follows the simulation, with the possible exception of a few solvation shells around the molecule itself. In this case a large fraction of the computational power is employed to the integration of the equations of motion

of degrees of freedom which are extremely relevant *during* the simulations, but are completely neglected *afterwards*.

In order to overcome this limitation, *coarse-grained models*[12–16] have been developed, where the structure and interactions of the original system are replaced with simpler ones, which are easier to describe, model, simulate and understand. The assumption underlying the coarse-graining of a system is that above a given length scale the low-level, chemistry-specific detail of the model affects some properties of the system only in a simple, functionally trivial way - often through prefactors. Examples of systems for which this approach proved to be extremely successful are molecular fluids, polymers[4,5], elastic network models of proteins[17–22], lipid membranes and other biomolecular systems, just to mention a few.

In recent years, systematic coarse-graining approaches have gained importance, where the interactions in the coarse-grained (CG) model are derived systematically from atomistic reference simulations in a bottom-up fashion. These models are often used in a multiscale simulation framework, where the closeness of higher and lower levels of resolution allows a switching back and forth between them. Below, we will review several systematic coarse-graining approaches and address some of the most important methodological issues and challenges.

The smaller number of degrees of freedom that are retained in coarse-grained models and the simpler force-fields employed allow to characterize relevant properties of a system at a cheaper computational cost compared to the high-resolution atomistic models; on the other hand, there are cases in which the chemical detail *in a small region of the system* plays a crucial role, such that no simplification of the description is possible: think, as an example, of the active site of a large enzyme, where fine-grained chemical processes take place. A high-resolution modeling of each part of the system would not be necessary, but at the same time a coarse-graining approach would delete important information.

This last observation naturally leads us to identify a particular class of soft matter systems among those that are studied with the help of computer simulations. Specifically, we can consider those systems where the focus is on a small, well-defined subregion of the simulation box. To this class belong, for example, certain solvated (macro)molecules, active sites of enzymes, the interaction of specific polymer ends at a surface, or simply a small spherical region in a homogeneous fluid whose radius is of the length scale of the property we're interested in.

For such systems the remaining, 'non-interesting' region is composed by the volume containing all those degrees of freedom which will be eventually neglected and/or discarded once the simulation is done, such as the solvent or large parts of a macromolecule which do not take active role in the process of interest (e.g. all atoms sufficiently far from the active site of an enzyme). Usually, the detailed knowledge about structural, energetic and thermodynamical properties of these large sections of the system is not required; nonetheless these 'non-interesting' degrees of freedom have to be explicitly present and integrated, inasmuch they 'scaffold' the target object of the simulation and represent a reservoir of energy and molecules.

A method is thus desirable, that allows one to perform a simulation where the largest part of the computational resources is concentrated on that region of the system that will be subsequently analyzed. *Adaptive resolution simulations methods*[23–33] were developed to solve the contradiction between the necessity of simulating all parts of the system and the

fact that, eventually, the detailed information referred to a large subgroup of them will be neglected. The underlying idea is to replace these 'non-interesting' degrees of freedom of the system with a simpler, coarse-grained representation, such that a sensibly smaller number of computations (e.g. force calculations) is required, while the 'interesting' region is treated at a higher resolution.

The present contribution discusses two strategies, the adaptive resolution simulation (AdResS) scheme and the Hamiltonian AdResS (H-AdResS) to perform simulations in which different regions of the same system are modeled with different resolution. Large parts of this text are based on/taken from a recent review[1] and some new developments coupling a liquid to an ideal gas[2], as well as previous work[34–37,29,32,33].

2 Coarse-Graining

Coarse-Grained models possess a number of features that make them particularly appealing. For example, a smaller amount of required computational resources due to both the reduced number of degrees of freedom and the simpler form of the interactions. Another important characteristic is that since many interaction centers are replaced with a single one, the fluctuations of the force experienced by a molecule are generally much smaller; this results in smoother free energy profiles and, as a consequence, in faster diffusive processes, allowing to reach larger time-scales with less computations. Finally, coarse-grained models are designed to entail large length-scale properties of the system, such as the global, collective conformational changes of a protein or the diffusive process of a polymer in a melt, that can be strongly insensitive to the fine-grained, chemistry-specific details; as a consequence, also the parametrization of the coarse-grained interactions is advantageously simpler.

Many CG models are generic, i.e. they were not developed to model a specific chemical system but rather with the aim to study a physical phenomenon such as polymer dynamics or folding or aggregation in general. One example are generic CG lipid models which have been successfully employed to study the self assembly of micelles, bilayers and other structures[38–42]. Generic CG models have also been employed to study folding and aggregation of peptides and proteins[43–55]. For polymers such generic models were especially successful. Following the so called 1/N theorem of de Gennes [56–58] it was shown that properties such as the overall chain extension as functions of the polymerisation index follow the same power law with the same exponent for all polymers, independent of the chemical species. The results of these scaling theories were instrumental for many developments also in connection to simulations of generic and thus very efficient models as well as to experiment. For dynamical properties generic simulations provided the first direct evidence of the reptation/tube concept put forward by Edwards and de Gennes [59,60]. The reptation model is based on the fact that the dynamics of long polymer chains is dominated by the constraint that polymer chains cannot simply cut through each other.
A wide range of approaches have been developed that aim for consistency between a CG model and either experimental data or simulations of accurate high resolution models. Typically, these approaches are divided into thermodynamics-based and so-called

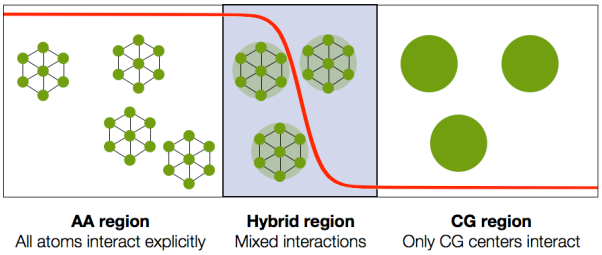


Figure 1: Typical scheme of an adaptive resolution simulation: a high-resolution region, where molecules are described at the atomistic level, is coupled to a low-resolution region where a simpler, coarse-grained model is employed. There two sub-parts of the system are interfaced *via* a hybrid region, in which the molecule's representation smoothly changes from one to the other, depending of their positions. It is on this last region and its properties (i.e. the way molecules change resolution) that the complexity of adaptive resolution schemes concentrates.

structure based ones. In thermodynamic coarse-graining approaches, individual elements of the CG interaction function are separately parameterized based on thermodynamic reference data such as solvation free energies and partitioning data, liquid densities, surface tension, etc.[61–72]. (These are usually experimental reference data, but in a multiscale simulation approach the reference data can of course also be obtained from an atomistic simulation, to keep the CG and atomistic level thermodynamically consistent). In another group of approaches one numerically generates CG interaction functions with the aim to reproduce the configurational phase space sampled in an atomistic reference simulation. These approaches may rely on different types of reference properties such as structure functions[73–85], mean forces[86–91] or relative entropies[92–94].

All CG models (in fact also all classical atomistic forcefields) are state-point dependent and cannot necessarily be – without reparametrization – transferred to different thermodynamic conditions or a different chemical environment compared to the one where they had been derived. This means 'transferability' can refer to a change in temperature, density, concentration, system composition, phase, etc., but also a change in chemical environment, e.g. the change of length or sequence of an amino acid chain. While there are many attempts to improve the transferability of coarse-grained models, there are situations, where one cannot easily refrain from considering more detailed interactions. In such situations adaptive resolution schemes are of advantage.

3 Adaptive Resolution Simulations

A typical system for which the focus of interest concentrates on a (possibly small) subregion of the simulated system is the case, for example, of the hydrogen bond network around a solvated molecule in water. The bulk of water molecules has to be simulated in order to sustain the thermodynamical properties of the subsystem of interest -the interfacial water- and to provide the correct exchange of molecules. Nonetheless, the fine-grained detail of molecules far from the interface is not relevant; it would be therefore desirable to replace the atomistic, expensive interactions of hydrogen and oxygen atoms with a coarser

model.
We can then introduce a geometrical separation between an 'inside' and an 'outside', i.e. an all-atom and a coarse-grained region, and assign different types of representations and interactions to the molecules according to their position in the simulation domain.
This idea has a long and successful history: to investigate crack propagation in hard matter, for example, several authors[95–99] made use of a hybrid description of the system, where a 'high resolution' description is employed only the area in proximity of the crack, and the material far from the latter is treated with a simpler model. Another important example of hybrid resolution simulation is provided by Quantum Mechanics / Molecular Mechanics (QM/MM) methods[100–104]. In this case the structure of the system is described at the same (atomistic) level everywhere; the interactions, though, are obtained from a classical force-field in the bulk of the system, but in a small region *ab initio* methods -such as Density Functional Theory, DFT- are employed to calculate the forces. Many different 'flavors' of this approach have been developed; in all of them, though, one of the crucial aspects is how to interface the two domains where interactions are different, and in most of the established methods the identity or resolution of the particle is not allowed to change. In general, one has to answer the two following questions:

1. how should two atoms/molecules in different domains interact?

2. how should the properties of an atom/molecule *change* in crossing the interface?

The last question is of particular importance for all systems whose components can diffuse on large length scales (at last of the order of the molecules' size) in the simulation time, i.e. essentially for almost all sooft matter systems. It appears natural to introduce a *transition region* (often called hybrid region, or healing region) that allows for a smooth interpolation from a given representation of the molecule's structure/interaction to another. The choice of the specific way this interpolation is implemented depends, as we mentioned earlier, on the properties that have to be preserved in the CG region.
Irrespective of the chosen method to interface the two regions of the system, though, it is natural to expect that the equilibrium state that will be reached in absence of external driving forces will not be the desired one. A further crucial point is then to find the simplest way to impose the desired thermodynamics.
The central, strong requirement that has to be satisfied is that molecules should be free to diffuse from any region of the simulation box to any other. Additionally, in a hybrid resolution model thermal equilibrium should be preserved, i.e., the temperature of the system has to be constant during the simulation. Another possible constraint is to impose a uniform density across the box, irrespective of the specific resolution; nonetheless, we'll see that there are cases where this is neither necessary nor desirable. Eventually this approach can be extended towards open systems MD.

3.1 The Adaptive Resolution Simulation scheme

The Adaptive Resolution Scheme (AdResS) represents the first, effective and computationally efficient method to simulate a system where two different models, e.g. an all-atom one and a coarse-grained one, are *simultaneously* employed in different subregions of the

simulation domain, interfaced in such a way to allow molecules to freely diffuse from one region to the other.
A basic constraint of the original version is that Newton's 3rd law has to be exactly satisfied everywhere. This requirement rules out any form of potential energy interpolation. It can be formally demonstrated[105] that no way exists to smoothly 'blend' the interaction between two molecules from a given potential energy to another without generating forces that cannot be recast in a form that satisfies Newton's Third Law. To preserve the latter, a *force-interpolation scheme* is required, which assures that forces are antisymmetric under exchange of the molecules' labels.

A second, less strict requirement is that CG molecules possess CG degrees of freedom only; this determines the specific way the force mixing is performed: a molecule in the CG region loses completely its atomistic detail (thus retaining, for example, the center of mass coordinates only), and interacts with a molecule in the AA or even the transition region only *via* its CG degrees of freedom. Formally, this constraint imposes that the atomistic forces vanish when at least one of the two interacting molecules is in the CG domain.
These two constraints are sufficient to define the force-field interpolation; the force acting between molecules α and β is given by:

$$\mathbf{F}_{\alpha\beta} = \lambda(\mathbf{R}_\alpha)\lambda(\mathbf{R}_\beta)\mathbf{F}^{AA}_{\alpha\beta} + (1 - \lambda(\mathbf{R}_\alpha)\lambda(\mathbf{R}_\beta))\,\mathbf{F}^{CG}_{\alpha\beta} \tag{1}$$

In Eq. 1 $\lambda(x)$ is any smooth function that goes from 1 in the AA region to 0 in the CG region. $\mathbf{R}_\alpha$ (resp. $\mathbf{R}_\beta$) is the CoM coordinate of molecule α (resp. β). $\mathbf{F}^{AA}_{\alpha\beta}$ and $\mathbf{F}^{CG}_{\alpha\beta}$ are, respectively, the atomistic and the coarse-grained forces acting on molecule α due to the interaction with molecule β. The CG force is computed between the coarse-grained centers of the molecules and then redistributed to the atoms weighted by the ratio of the atom's mass to the mass of molecule[106]; in the transition region this operation is required by the fact that molecules interact at both the AA and the CG level. AA degrees of freedom thus *have* to be explicitly integrated, at least into the hybrid region. In the CG region, on the other hand, it is in principle not necessary to conserve the atomistic detail of the molecules, so that the CG force could be applied directly to the CoM coordinate; a molecule's internal structure can thus be removed when it enters the CG region, and reintroduced (e.g. taking it from a reservoir / repertoire of equilibrated atomistic molecules) as soon as it approaches the hybrid region.

Since the force interpolation based AdResS scheme cannot be formulated in terms of a Hamiltonian, it is impossible to perform microcanonical, i.e. energy-conserving simulations. The force-field used in this adaptive resolution simulation framework is not conservative in the transition region. This excess energy of the transition region can be removed with a *local* thermostat, such as Langevin thermostat. The equilibrium state of the system is then *dynamical*: the thermostat takes care of absorbing the extra heat produced in the transition region by non-conservative forces, and the system samples equilibrium configurations according to Boltzmann's distribution[23–31].
The different pressure between an AA system and a low-resolution model typically resulting from coarse-graining procedures determines the onset of a non-uniform density profile. We already mentioned, for example, that a one-site CG model of water obtained with IBI can have a pressure ~ 6000 times the atomistic reference value[107]. Therefore, the densities

in the two subregions will change in order to equate the pressures. Correcting the pressure would result in a modified compressibility[107], which is not desirable in most cases. An option to preserve a uniform density across the simulation domain without modifying the CG potential is to introduce an external force which counterbalances the high pressure of the CG model. This *thermodynamic force* can be obtained with an iterative procedure *via* the following expression[108,106,109]:

$$\mathbf{f}_{th}^{i+1} = \mathbf{f}_{th}^{i} - \frac{1}{\rho^{\star}\kappa_T}\nabla\rho^{i}(r) \tag{2}$$

where $\rho^{\star}$ is the reference molecular density, κ_T is the system's isothermal compressibility and $\rho^{i}(r)$ is the molecular density profile as a function of the position in the direction perpendicular to the CG-AA interface. The thermodynamic force is initialized to zero, $\mathbf{f}_{th}^{0} = 0$, while the initial density profile is the one calculated from an AdResS simulation with $\mathbf{f}_{th} = 0$. As it can be easily seen, the iterative procedure converges once the density profile is flat ($\nabla\rho(r) = 0$). This approach guarantees a flat density profile without having to modify the CG potential.

In summary, the thermodynamic force allows us to couple a system at atomistic resolution to a coarse-grained counterpart whose pressure, for given values of density and temperature, is sensibly different. The global properties of the force, whose direct effect is restricted to the hybrid region, only depend on the pressure difference between the two coupled subsystems; the detailed profile of the force, on the other hand, can be obtained *via* a system-specific iterative procedure. This method not only allows to preserve the desired structure of the system in the CG region; in principle, in fact, an arbitrary CG force-field, with pressure *and structure* completely off from the atomistic target ones, can be used. Consequently, the AA region behaves as an open system[106] that exchanges energy and molecules with a reservoir: the molecule number fluctuations, the pressure and all other thermodynamically relevant quantities are the same as if the AA region were simply 'cut' from a large all-atom simulations. It is relevant to stress here that because of the thermodynamic force this condition can be established *irrespective of the specific model used in the CG region*.

3.2 Applications

The possibility of treating a system with a reduced number of degrees of freedom except where it is strictly necessary was explored, making use of the AdResS method, in several applications[23–26,110,27–29]. From the numerical/computational point of view it clearly represents an advantage, since a much smaller number of force calculations are required in the coarse-grained region: this is particularly true for parallel MD codes such as GROMACS[111], where a dynamical decomposition of the simulation box allows one to subdivide the latter with a finer grid in the AA and hybrid region, while a smaller number of processors is assigned to the CG region. For example, for a water system with an AA region covering $1/6$ of the total simulation box, simulated with GROMACS on a 16-cores processor, the speed-up is about a factor three. This factor is nonetheless small compared to what can be achieved with other simulation packages, such as ESPRESSO++[112]: in fact, water simulation in GROMACS is extremely optimized, and any hacking of the

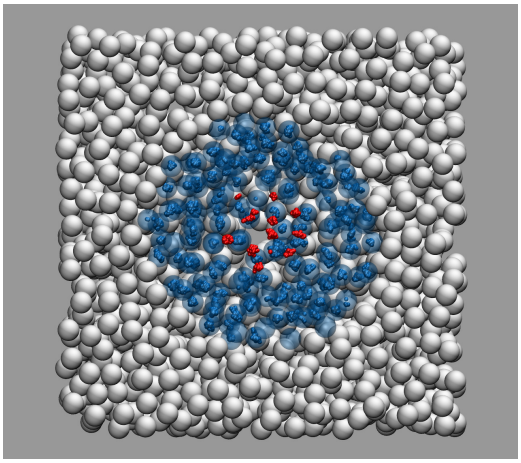


Figure 2: Set-up of the AdResS para-hydrogen simulation performed in Ref.[29] (figure adapted from therein). A small sphere in the center of the box, having radius as small as 0.6 nm, is treated at the path integral level (red rings), while the rest is described by point-like molecules (the white spheres); the hybrid region (blue) interfaces these two representations.

standard code can introduce a bottleneck.

This method has already been applied to several problems, like the work in Ref.[113]: here a molecule with both hydrophilic and hydrophobic interactions was solvated in water and put at the center of the high-resolution region, while the water molecules far from the surface were treated at the coarse-grained level. The ordering degree of the hydrogen bond network on the molecule's surface was measured as a function of the size of the all-atom region: the results showed a dependency of the ordering for water molecules close to the surface of the repulsive solute, while no relevant effect was observed for the attractive case. The same strategy has been applied to investigate the extent of spatial correlations in a quantum fluid, namely low-temperature para-hydrogen[114,29]. The latter is the spin-zero singlet state of molecular hydrogen. Because of the spherical symmetry of the global wave function, para-hydrogen in the solid and gas phase can be modeled as a classical, point-like particle interacting *via* a simple radial potential, such as Lennard-Jones or the more accurate Silvera-Goldman potential[115,116]. The possibility to simulate a quantum system in a classical framework such as classical MD makes it possible to couple quantum an classical descriptions with the AdResS scheme.

More recently the AdResS scheme has been successfully employed to perform simulations of biologically relevant systems such as methanol-water mixtures[117] and triglycine in aqueous urea[109], and to study the coil-globule transition of a PNIPAm molecule in aqueous methanol[118]. In all these cases a crucial necessity is to correctly reproduce the solvation free energies of the system, a condition that is verified only when the particle number fluctuations are compatible with those observed in the Grand Canonical ensemble. The large system sizes necessary to fulfill this requirement in a standard, all atom simulation often make the latter unfeasible; the employment of dual-resolution simulation methods, pos-

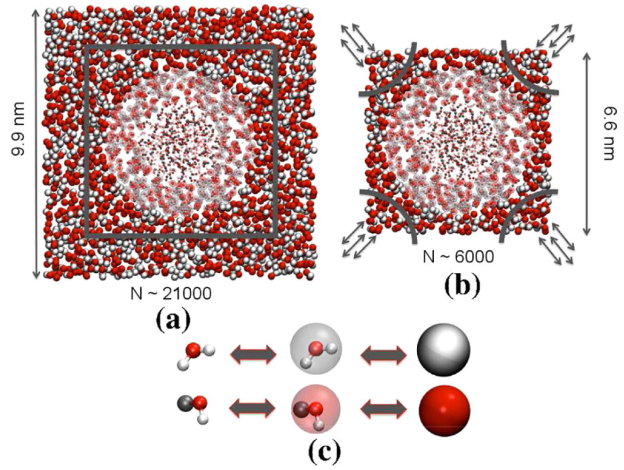


Figure 3: Schematic representation of the schemes used for the simulations of a PNIPAm molecule solvated in aqueous methanol: (a) Conventional AdResS scheme, where a small AA region is coupled to a large "closed boundary" coarse-grained reservoir. (b) Particle exchange adaptive resolution scheme (PE-AdResS), where an AA region is coupled to a much smaller open boundary coarse-grained reservoir, where particle exchange is performed at the eight corners of the simulation domain to avoid depletion effects. (c) Mapping scheme representing the smooth coupling between AA and CG particle representations. Figure from[118].

sibly coupled to a Monte Carlo scheme[118] to enforce fluctuations in the total number of molecules, see fig. 3, allows one to keep the computational cost low and obtain results that would otherwise require a sensibly longer time.

3.3 The limitations of the force-based approach

The AdResS method discussed so far represents a simple, effective way to perform double-resolution simulations with well controlled thermodynamics, i.e. simulations where the model used to represent a molecule and its interactions with the others change according to the molecule position. Furthermore the external field -the thermodynamic force- in the hybrid region allows one to *couple arbitrarily different systems* while keeping locally well-defined temperature, pressure and energy.

The AdResS method was conceived based on the requirement that Newton's Third Law has to be exactly satisfied everywhere. *A posteriori*, the resulting lack of a global energy function proves not to be a major problem: equilibrium and canonical sampling can be enforced making use of a local Langevin thermostat. A theoretical analysis of the AdResS dual resolution scheme has been recently carried out in Ref.[119], where the presence of a local thermostat and the thermodynamic force have been shown to be necessary and sufficient conditions to guarantee the equivalence of the atomistic region to an open region of a fully atomistic simulation up to second order correlation functions (density profile and radial distribution function).

Nonetheless, the lack of a Hamiltonian restricts the use of the AdResS method: micro-canonical, i.e. energy-conserving simulation are not possible; there is no partition function for the system as a whole; no Monte Carlo scheme can be implemented, and the system has

to be locally thermostatted in the hybrid region[31]. In the next section a method is discussed, named H-AdResS[32] (for Hamiltonian Adaptive Resolution Simulation scheme), that provides a solution to the aforementioned problems; clearly, as there is no free lunch, there is a price to pay: the Hamiltonian formulation requires a local breakdown of Newton's Third Law.

3.4 The Hamiltonian Adaptive Resolution scheme: H-AdResS

As was discussed in the previous section, if a position-dependent interpolation of the potential energies is performed, the resulting forces include a term proportional to the derivatives of the switching function λ that cannot be recast in a form that satisfies Newton's Third Law. The only method developed so far that allows one to explicitly conserve the energy in an adaptive resolution simulation is the one proposed by Heyden and Truhlar,[120,121], where a sum of the Lagrangians of all possible grouping of atomistic and coarse-grained molecules is done. Due to its combinatoric nature, this approach is extremely difficult to implement efficiently; moreover, the resulting Lagrangian includes a position-dependent kinetic energy term for which a specific, non-symplectic integrator is required.
In the H-AdResS method[32], which we now describe, the aforementioned constraints are relaxed in order to develop an energy-based, Hamiltonian adaptive resolution simulation scheme. As will be clear in a few lines, the particular choice of energy 'mixing' gives rise to forces that do not comply with the first constraint; nevertheless, the physical interpretation of these terms is immediate and naturally points towards the solution -though approximate- of Newton's Third Law breakdown.
The core idea of the energy-based approach is to weight the *total energy* of each molecule with a position-dependent function:

$$H = \mathcal{K} + V^{int} + \sum_{\alpha} \left\{ \lambda_\alpha V_\alpha^{AA} + (1-\lambda_\alpha) V_\alpha^{CG} \right\} \tag{3}$$

where $\mathcal{K}$ is the (all-atom) kinetic energy of the molecules, V^{int} is the interaction internal to the molecules, and:

$$\begin{cases} V_\alpha^{\mathrm{AA}} \equiv \dfrac{1}{2} \displaystyle\sum_{\beta,\beta\neq\alpha}^{N} \sum_{ij} V^{AA}(|\mathbf{r}_{\alpha i} - \mathbf{r}_{\beta j}|) \\ V_\alpha^{\mathrm{CG}} \equiv \dfrac{1}{2} \displaystyle\sum_{\beta,\beta\neq\alpha}^{N} V^{CG}(|\mathbf{R}_\alpha - \mathbf{R}_\beta|) \\ \lambda_\alpha = \lambda(\mathbf{R}_\alpha) \end{cases}$$

The switching function λ goes from 0 (purely CG) to 1 (purely AA). The force acting on atom i in molecule α is obtained through differentiation of the Hamiltonian in Eq. 3:

$$\begin{aligned} \mathbf{F}_{\alpha i} = & \mathbf{F}_{\alpha i}^{int} \\ & + \sum_{\beta,\beta\neq\alpha} \left\{ \frac{\lambda_\alpha + \lambda_\beta}{2} \mathbf{F}_{\alpha i|\beta}^{AA} + \left(1 - \frac{\lambda_\alpha + \lambda_\beta}{2} \right) \mathbf{F}_{\alpha i|\beta}^{CG} \right\} \\ & - \left[V_\alpha^{AA} - V_\alpha^{CG} \right] \nabla_{\alpha i} \lambda_\alpha \end{aligned} \tag{4}$$

The forces $\mathbf{F}^{AA}_{\alpha i|\beta}$ and $\mathbf{F}^{CG}_{\alpha i|\beta}$ are defined as:

$$\mathbf{F}^{AA}_{\alpha i|\beta} \equiv \sum_{j=1}^{n_\beta} -\frac{\partial}{\partial \mathbf{r}_{\alpha i}} V(|\mathbf{r}_{\alpha i} - \mathbf{r}_{\beta j}|)$$
$$\mathbf{F}^{CG}_{\alpha i|\beta} \equiv -\frac{m_{\alpha i}}{M_\alpha} \frac{\partial}{\partial \mathbf{R}_\alpha} V^{CG}(|\mathbf{R}_\alpha - \mathbf{R}_\beta|) \tag{5}$$

The redistribution of the CG force on the atomistic degrees of freedom follows the same rules as applied in the case of the force-based AdResS method. It's worth noting that in this energy-based scheme the atomistic degrees of freedom are retained and integrated everywhere in the system, a necessary requirement in order to perform a *microcanonical* simulation making use of a Hamiltonian.

The first term, $\mathbf{F}^{int}_{\alpha i}$, is due to internal interactions of the molecule; as such, it automatically satisfies Newton's Third Law. The second term is a sum of pairwise forces obtained from all-atom and coarse-grained Hamiltonians, weighted by a function that is symmetric under molecule label exchange, that is $\alpha \leftrightarrow \beta$; also this force complies with Newton's Law. The third term of the forces in Eq. 4 is the part that breaks Newton's Third Law. This force, that is nonzero *only in the hybrid region*, is proportional to the difference between the potential energies of a given molecule in the AA and the CG representation; if a systematic difference exists between the AA and the CG potentials, the effect of this term is to push molecules in one of the two bulk regions. The hybrid region thus behaves as an *active membrane*, inducing a density imbalance and a non-flat pressure profile. Ideally, *when the CG potential perfectly reproduces the many-body PMF*, this drift would vanish on average:

$$V^{CG}_{\alpha\beta} \equiv \left\langle V^{AA}_{\alpha\beta} \right\rangle \;\Rightarrow\; \langle \mathbf{F}^{dr}_\alpha \rangle \propto \left\langle \left[V^{AA}_\alpha - V^{CG}_\alpha \right] \right\rangle \to 0$$

Needless to say, the CG potentials almost never reproduce the many-body potential of

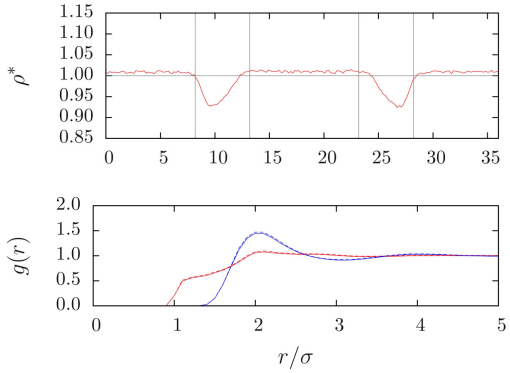


Figure 4: H-AdResS simulation of a system of tetrahedral molecules coupled to point-like molecules interacting through an IBI-CG potential. Top: density profile; bottom: radial distribution functions of the atomistic (red lines) and coarse-grained (blue lines) degrees of freedom in the all-atom region; the solid lines are the reference RDF's calculated in the all-atom system, while the dashed lines are obtained from a H-AdResS simulation.

mean force[122,107]. The difference between an atomistic model and its coarse-grained representation therefore results in a thermodynamic imbalance, that is, both pressure and density of the two bulk (AA and CG) regions are different[14]. The solution to this problem is again to introduce a compensation term in the Hamiltonian, as it was done in the AdResS scheme with the thermodynamic force. More specifically, we modify the Hamiltonian as follows:

$$H_\Delta = H - \sum_{\alpha=1}^{N} \Delta H(\lambda(\mathbf{R}_\alpha)) \tag{6}$$

where $\Delta H(\lambda)$ is a function to be defined. It's worth noting that this term preserves the conservative nature of the Hamiltonian.
In order to determine the specific form of ΔH we impose that the drift force cancels on average:

$$\left.\frac{d\Delta H(\lambda)}{d\lambda}\right|_{\lambda_\alpha} \nabla_\alpha \lambda_\alpha + \langle \mathbf{F}_\alpha^{dr} \rangle \equiv 0 \tag{7}$$

or equivalently:

$$\left.\frac{d\Delta H(\lambda)}{d\lambda}\right|_{\lambda=\lambda_\alpha} = \left\langle \left[V_\alpha^{AA} - V_\alpha^{CG}\right] \right\rangle_{\mathbf{R}_\alpha} \tag{8}$$

where the subscript in the average indicates that the latter has to be performed constraining the CG site of molecule α in the position $\mathbf{R}_\alpha$.
In principle, Eq. 8 provides us with the way to compute the compensating function - or, more precisely, its derivative; nonetheless, an approximation to ΔH might be sufficient. A way to do this is the following:

$$\left\langle \left[V_\alpha^{AA} - V_\alpha^{CG}\right] \right\rangle_{\mathbf{R}_\alpha} \simeq \frac{1}{N} \left\langle \left[V^{AA} - V^{CG}\right] \right\rangle_{\lambda'} \tag{9}$$

where $\lambda' \equiv \lambda(\mathbf{R}_\alpha)$ is the same for all molecules. The approximate function ΔH is obtained by integration:

$$\begin{aligned} \Delta H(\lambda) &= \int_0^\lambda d\lambda' \frac{d\Delta H(\lambda')}{d\lambda'} \\ &\simeq \frac{1}{N} \int_0^\lambda d\lambda' \left\langle \left[V^{AA} - V^{CG}\right] \right\rangle_{\lambda'} = \frac{\Delta F(\lambda)}{N} \end{aligned} \tag{10}$$

Most interestingly the compensation needed to cancel $\langle \mathbf{F}_\alpha^{dr} \rangle$ is related to the *Helmholtz free energy* difference between AA and CG system[123]. Therefore, it is possible to calculate the compensating function needed to restore, on average, Newton's Third Law by performing a Kirkwood thermodynamic integration.
The 'Helmholtz free energy compensation' thus cancels the active effect of the hybrid region, restoring a flat pressure profile. Nonetheless, coarse-grained models have, in general, a substantially different pressure with respect to their atomistic reference[107], thus inducing a further density imbalance (usually larger than the one due to the different Helmholtz free energy). In order to restore a flat density profile a second term has then to be added to the compensating function, that counterbalances the pressure difference.
The right way to introduce the pressure in the compensating function is to balance, rather

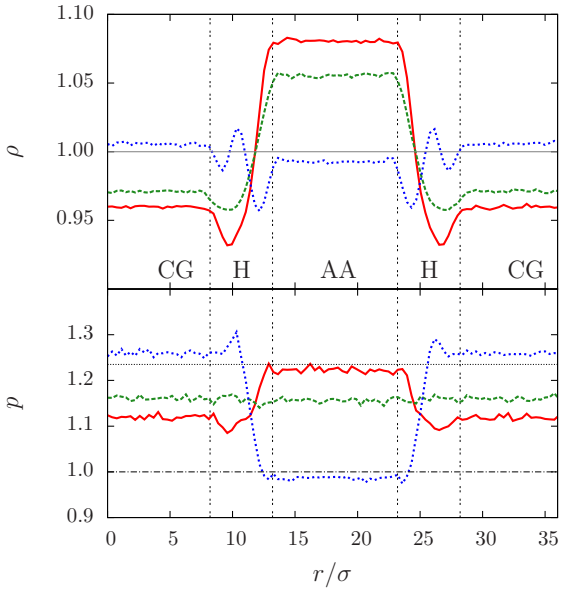


Figure 5: Plots showing the effect of the free energy compensations on the density profile (upper panel) and pressure profile (lower panel) in a H-AdResS simulation with CG potential having larger pressure, for identical temperature and density values, than the all-atom one. The red line corresponds to the case where no compensating function was employed; the green lines to the Helmholtz free energy compensation; and the blue line to the Gibbs free energy compensation. All densities are normalized to the value of the fully atomistic simulation (dotted line at $\rho = 1$). All pressures are normalized to the value of the fully atomistic simulation (trait-dot line); the dotted line indicates the normalized pressure of the fully coarse-grained simulation.

than Helmholtz free energy, the *Gibbs free energy* difference per particle, that is, the chemical potential $\Delta\mu = \Delta G/N$:

$$\Delta H(\lambda) \equiv \Delta\mu(\lambda) = \frac{\Delta F(\lambda)}{N} + \frac{\Delta p(\lambda)}{\rho^\star} \tag{11}$$

Fig. 5 shows the density and pressure profiles for the three possible cases discussed above. Analogous results are obtained in a thermostatted simulation of a water box, as shown in Fig. 6: here the system is composed by a slab of water molecules described at atomistic resolution, coupled to a CG bulk where particles interact *via* a purely repulsive WCA potential. As in the previous case, the CG interaction was parametrized to induce an increase of the density in the atomistic region, as can be seen in Fig. 7 (upper panel). The Free Energy Compensation restores the correct density profile, and guarantees that in the AA region the pairwise correlations, i.e. the radial distribution functions, are the same that one would measure in a fully atomistic simulation, as shown in Fig. 7 (bottom panel). We notice that Gibbs free energy compensation, even though it equates the densities in the bulk regions, is not sufficient to remove small fluctuations (of the order of $\sim 3\%$) in the hybrid region: these deviations from the reference value are due to the fact that the compensation ΔH is computed in a homogeneous system, where all molecules have the same value of λ - that is, a regular Kirkwood thermodynamic integration Hamiltonian. The molecules

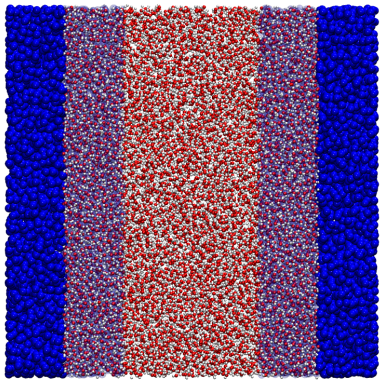


Figure 6: Schematic view of a dual-resolution simulation of water: the central slab of the box is described at atomistic resolution, while in the bulk the molecules are point-like particles interacting *via* a purely repulsive WCA potential.

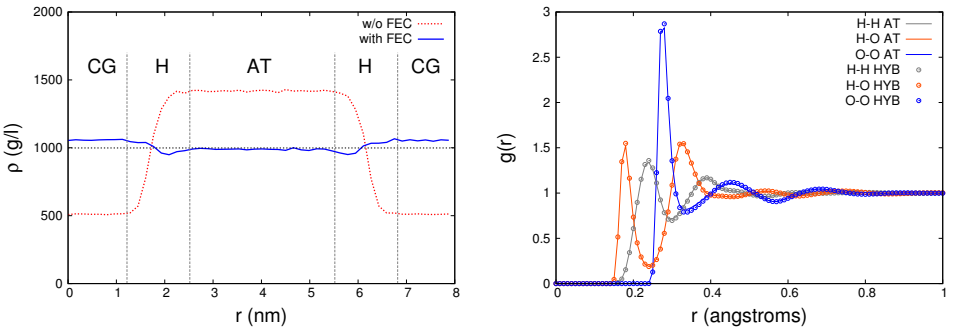


Figure 7: Left panel: density profile of the water system along the x coordinate. The red dotted line corresponds to the H-AdResS simulation without FEC, while the solid back line has been obtained making use of the FEC. Right panel: radial distribution functions of the water atoms in the central (AT) slab of the box, as obtained from a fully atomistic simulation (solid lines) and a H-AdResS simulation with FEC (dots).

in the hybrid region, on the other hand, interact with other molecules having different λ values. The resulting fluctuations are expected to decrease with increasing size of the hybrid region, in which case the environment of a given molecule approaches the condition of homogeneous λ. Another strategy to flatten the density profile is clearly provided by the iterative approach of the thermodynamic force (Eq. 2), a few iterations of which would be sufficient to modify the ΔH function by the small amount necessary to remove the fluctuations.

The Free Energy Compensation (FEC) strategy, defined by Eq. 6, can be extended to multi-component systems. To illustrate this idea we consider a molecular liquid composed by two types of molecules, A and B, indexed with a and b, respectively. The corresponding

H-AdResS Hamiltonian for this system reads:

$$H^{MIX} = K + V^{int} + \sum_{a\in A}\left[\lambda_a V_a^{AA} + (1-\lambda_a)V_a^{CG}\right] + \sum_{b\in B}\left[\lambda_b V_b^{AA} + (1-\lambda_b)V_b^{CG}\right] \tag{12}$$

with $\lambda_a = \lambda(\mathbf{R}_a)$ and $\lambda_b = \lambda(\mathbf{R}_b)$. The intermolecular potential energy terms are given by the following expressions:

$$\begin{aligned}
V_a^{AA} &= \frac{1}{2}\left[\sum_{\substack{a'\in A\\ a'\neq a}}\sum_{ij} V[AA]^{AA}_{ai;a'j} + \sum_{b\in B}\sum_{ij} V[AB]^{AA}_{ai;bj}\right]\\
V_a^{CG} &= \frac{1}{2}\left[\sum_{\substack{a'\in A\\ a'\neq a}} V[AA]^{CG}_{aa'} + \sum_{b\in B} V[AB]^{CG}_{ab}\right]\\
V_b^{AA} &= \frac{1}{2}\left[\sum_{\substack{b'\in B\\ b'\neq b}}\sum_{ij} V[BB]^{AA}_{bi;b'j} + \sum_{a\in A}\sum_{ij} V[AB]^{AA}_{bi;aj}\right]\\
V_b^{CG} &= \frac{1}{2}\left[\sum_{\substack{b'\in B\\ b'\neq b}} V[BB]^{CG}_{bb'} + \sum_{a\in A} V[AB]^{CG}_{ba}\right]
\end{aligned} \tag{13}$$

where $V[XY]$ is the non-bonded interaction between a molecule of type X and a molecule of type Y, with $X, Y = A, B$, and the indices i, j label the atoms.

In analogy with one-component systems we introduce a FEC term for each species to compensate for the free energy difference between the AA and the CG regions:

$$H_\Delta^{MIX} = H^{MIX} - \sum_{a\in A}\Delta H_A(\lambda_a) - \sum_{b\in B}\Delta H_B(\lambda_b) \tag{14}$$

An *Ansatz* for the compensation term of a given species $k = a, b$ can be obtained from TI as follows:

$$\begin{aligned}
\Delta H_k(\lambda) &= \frac{\Delta F_k(\lambda)}{N_k} + \frac{\Delta p_k(\lambda)}{\rho_k^\star}\\
\Delta F_k(\lambda) &= \int_0^\lambda d\lambda' \left\langle \left[V_k^{AA} - V_k^{CG}\right]\right\rangle_{\lambda'}\\
\Delta p_k(\lambda) &= p_k(\lambda) - p_k(0)
\end{aligned} \tag{15}$$

where the N_k, $\rho_k^\star \equiv N_k/V$ and p_k are, respectively, the number of molecules, the reference partial density and the partial virial pressure of species k. We stress that all the quantities in Eq. 15 can be computed in a single TI of the mixture from AA to CG at the concentration of interest, irrespective of the number of species. All the cross-interactions between different types of molecules are automatically included in the free energy contribution of each

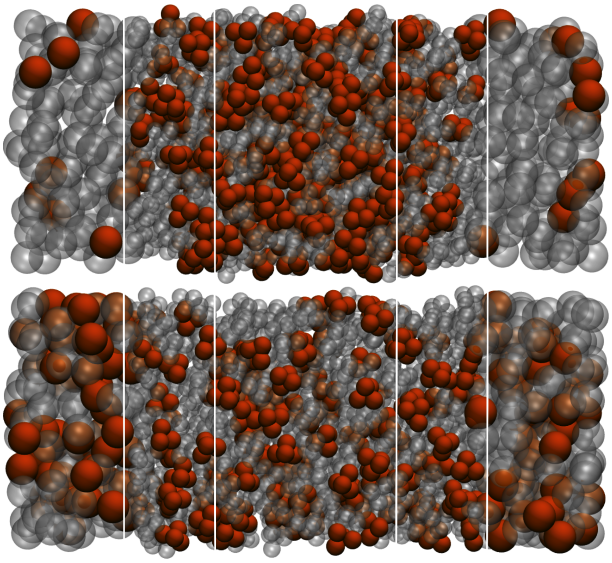


Figure 8: Snapshots of a H-AdResS Monte Carlo simulation. Top panel: equilibrated configuration, without FEC. Bottom panel: equilibrated configuration, with FEC. The A-type atoms are represented in gray, the B-type atoms in orange. Molecules in the coarse-grained (CG) region are represented as large spheres. White vertical lines mark the boundaries of the CG-hybrid and hybrid-atomistic regions.

species. Additionally, the Free Energy Compensation $\Delta H_k(\lambda)$ is an *intensive* quantity and does not depend on the specific geometry of the H-AdResS setup. It is therefore possible to perform the TI in a relatively small system, provided that it is statistically representative, *i.e.* finite size effects are negligible.

The effectiveness of this strategy has been proven by the Monte Carlo simulations of binary mixtures performed in Ref.[33]. Here we report one of these simulations, specifically the mixture of 70% A-type molecules and 30% B-type molecules, both made of four identical atoms; the A–A and B–B interactions are identical WCA potentials, while the A–B interaction is a Lennard-Jones potential. In the CG region both molecules are represented as spherical particles with identical, purely repulsive WCA A–A, B–B and A–B interactions, resulting in a particularly large thermodynamic mismatch between AA and CG domains. This can be directly observed in the snapshot of the simulation reported in Fig. 8 (top) as well as in the density profiles (dotted lines in Fig. 9): the chemical potential imbalance between the two resolutions determines a large accumulation of B-molecules in the AA zone. As a consequence, neither the total density nor the relative concentrations in the AA zone obtained using the uncompensated adaptive resolution Hamiltonian in Eq. 12 correspond to the reference atomistic system.

3.5 Coupling a liquid to an ideal gas

In most cases, one tries to preserve, in the coarse-grained model, certain fundamental properties of the higher resolution system. The fact that one can essentially couple arbitrary systems has recently been demonstrated for both AdResS and H-AdResS by

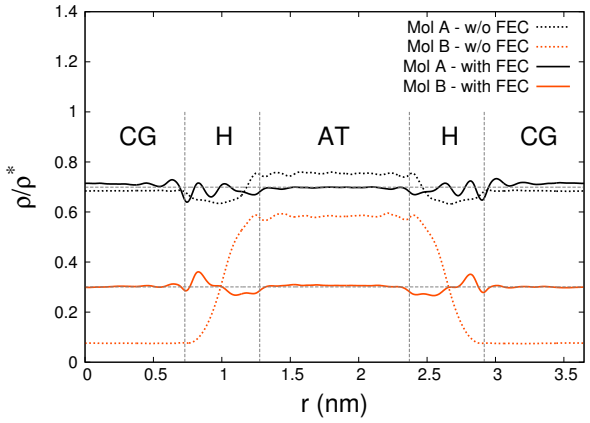


Figure 9: Density profiles along the direction of resolution change. Dotted lines: H-AdResS simulations without FEC; solid lines: with FEC. Vertical dashed lines indicate the boundaries between the AT, hybrid and CG regions; horizontal dashed lines mark the reference value of the density (normalized to the total density) as expected in a fully atomistic simulation of the system.

performing adaptive resolution simulations of water coupled to an ideal gas[2]. It was shown that the density as well as the liquid structure in the atomistic domain was not altered by the ideal gas environment.

In fact, if the scope of the simulation does not focus on the low resolution part of the system, the choice of the coarse-grained interaction may be driven mainly by system simplicity and computational efficiency, provided that the correct thermodynamical conditions are preserved in the high resolution domain. In this case, it is easy to see that the optimal choice is to have no interactions in the low resolution domain at all. In practical terms, this means coupling the system in the high resolution region to an ideal gas. Another advantage is the possibility of varying the number of particles in the system at will by simply creating or destroying molecules in a region of the system where they are all "invisible" to each other. This would straightforwardly allow the simulation of a system with fixed chemical potential rather than number of particles, that is, to simulate a truly Grand Canonical ensemble with minimal computational effort. This approach has recently been followed to simulate liquid water. The large computational benefit based on the simplicity of an ideal gas comes at the cost of large thermodynamic differences between the equations of state of an ideal gas and a molecular fluid. These free energy discrepancies nontrivially affect the behavior of the hybrid, dual resolution system, and require particular care in the construction of the interface. Several strategies have been developed to provide the thermodynamic balance between the two coupled models[106,32], and have been discussed above. Still, the large free energy difference, as well as diffusion dynamics that differ by orders of magnitude, make the construction of a smooth seam between water and ideal gas a challenging problem as illustrated in Fig. 10.

In Fig. 11 we report the density profiles measured in the H-AdResS simulations of an atomistic water model coupled to the ideal gas (the details about the simulations can

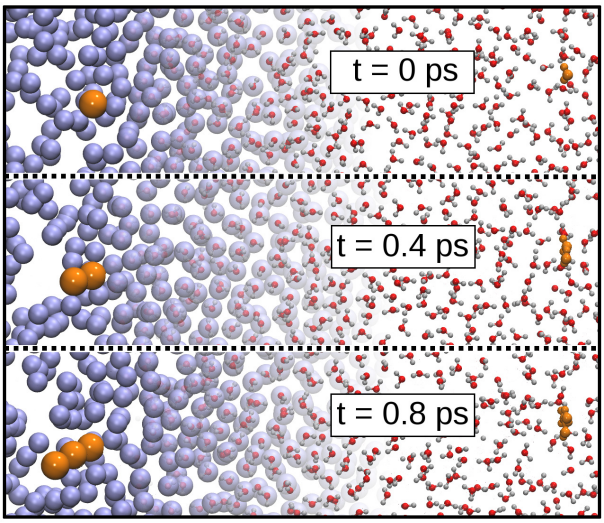


Figure 10: Three consecutive snapshots of the water–ideal gas interface. For clarity we show only molecules in a 0.7 nm thick layer in the direction perpendicular to the sheet. The water, in the right half of the figure, is much more structured than the ideal gas (left), in which several molecules overlap. For the orange highlighted molecules the time evolution is visualized by copying their previous positions into the subsequent snapshots. From this it can be seen that the molecules in the ideal gas diffuse much faster than the water molecules, whose positions almost do not change over 0.8 ps.

be found in[2]). A simple coupling of these models leads to a density imbalance, due to their different equations of state. It is possible, however, to remove such mismatch by introducing a free energy compensation term. This approach leads to an almost flat density profile, with the exception of a small deviation from the reference in the interface between hybrid and coarse-grained regions. What is relevant, though, is that the density attains the correct reference value in the atomistic region, where also the atomistic water structure is unperturbed by ideal gas surrounding, as it can be seen from the RDFs in Fig. 12.

A crucial point is to verify that the system in the AT region behaves as if it were a subpart of a completely atomistic system. This means not only measuring the density profile and the RDFs, but also checking that the molecules do not have any impediment in diffusing across the HY region. To this end, we followed the time evolution of a subset of labelled particles at the AT/HY and CG/HY interface, to check that nothing prevents them from moving across the system. Obviously, we can expect a different diffusion rate in the two cases, as the friction of the ideal gas is decidedly smaller than that experienced by the atomistic water molecules.

In Fig. 13 (solid lines) we report the diffusion profiles of the molecules initially located in two symmetric slabs of width 1 nm in the AT region, at the interfaces with the HY regions. These molecules, uniformly distributed at $t = 0$, spread out throughout the whole system as time passes. The overall distribution is quasi-Gaussian, but the half moving towards the CG region extends further than the half in the AT region, as the friction in the former allows a faster diffusion. The diffusion in the AT region, though, is perfectly compatible with that of a fully atomistic water system. This can be verified by comparing

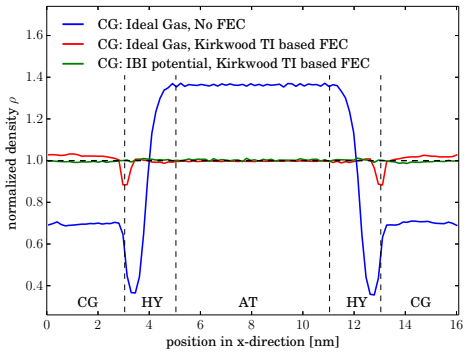


Figure 11: Density profiles for H-AdResS simulations of the water–ideal gas system with and without Kirkwood TI based Gibbs FEC as well as for the water–IBI system, also with Kirkwood TI based Gibbs FEC.

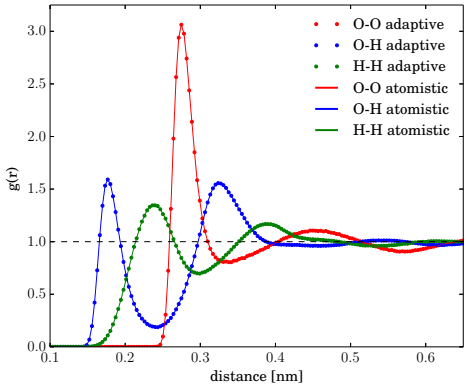


Figure 12: RDFs for pure water and for the atomistic region of the H-AdResS water–ideal gas with Kirkwood TI based Gibbs FEC. Since a rigid water model is employed, all RDFs consider only intermolecular atom pairs.

the diffusion profiles of the H-AdResS simulation with those obtained by performing the same analysis on a fully atomistic system. These latter distributions, reported in Fig. 13 (dashed lines), overlap very well in the AT region and even in part of the HY region, while, as expected, the distribution of the molecules in the H-AdResS simulation (solid lines) extends deeper in the CG region.

A fundamental measure of the system behavior is provided by the local fluctuations in the particle number, defined as:

$$\Delta N \equiv \frac{\langle N^2 \rangle - \langle N \rangle^2}{\langle N \rangle} \tag{16}$$

where N is the number of particles in a 1-nm wide slab of the simulation box. Preserving the correct number fluctuations in the high resolution region implies that the solvation properties in the latter subsystem are representative of a fully atomistic simulation.

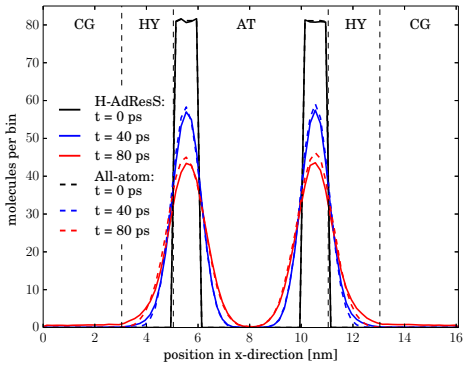


Figure 13: Diffusion profiles in H-AdResS simulations of the water-ideal gas system and in fully atomistic reference simulations of SPC/E water: time evolution of the position of molecules initially located in a 1-nm-wide slab in the atomistic region, immediately adjacent to the HY region. The y-axis is the absolute number of these molecules whose centre of mass X-coordinate is in a given bin at the given time.

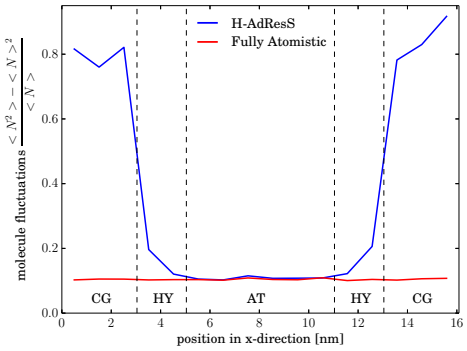


Figure 14: Molecule fluctuations as a function of position in the simulation box for H-AdResS water–ideal gas simulation with Gibbs FEC as well as for fully atomistic reference simulations.

The resulting profiles are shown in Fig. 14. Although the local density fluctuations in the ideal gas region are almost an order of magnitude higher than those in the atomistic region, the latter nonetheless correspond perfectly to the local density fluctuations in a fully atomistic system.

4 Conclusions

The characterization of the properties of new materials, as well as the investigation of biological macromolecular machineries, have largely benefited from *in silico* experiments. In spite of a steady increase in the available computational power, though, for very large systems and long timescales of the precesses involved these resources turn out to be insufficient, due to the extraordinarily large amount of data that has to be stored and force/energy calculations that have to be performed. To overcome these limitations,

the field of multiscale simulations has vastly expanded over the last years, and in the present review we have covered two aspects that are central to many multiscale approaches.

Here we have concentrated on recent advances in the field of adaptive resolution approaches. The above mentioned limitation in system size comes together with the fact that a considerable fraction of the simulated data is often discarded afterwards. Adaptive resolution methods try to reduce the amount of resources dedicated to the simulation of large, non-interesting regions of the system by replacing them with a simpler, coarse-grained representation of their content. Such 'dual-resolution' schemes are built with the constraint that the thermodynamic properties of the region of interest (i.e. the one with the higher resolution) do not differ from those that an equivalent subdomain of the system would have in a fully high-resolution simulation.
In the present work we discussed two methods to achieve this goal: the Adaptive Resolution Simulation (AdResS) scheme, based on the interpolation of two different force-fields, and its Hamiltonian formulation, H-AdResS, where the all-atom and coarse-grained potential energies are interpolated. These methods have been successfully applied to interface different molecular fluids, treated at the atomistic level, with their coarse-grained models; the different properties of the AA and the CG potentials naturally induce thermodynamical imbalances in the corresponding sub-regions, but simple and effective ways to overcome this problem have been described.
The possibility to replace vast regions of the simulated system with a crude, computationally inexpensive representation -even an ideal gas- and concentrate the computational resources on smaller parts while keeping the relative thermodynamics under control makes it possible to sensibly reduce the amount of calculations required to perform a simulation, and opens the way to a broad spectrum of applications, such as large-scale simulations of complex biomolecules in solution and efficient open-boundary simulations with varying number of particles.

Acknowledgments

We thank all present and past members of the multiscale modeling group at the Max Planck Institute for Polymer Research, the theoretical chemistry group at the University of Konstanz as well many further colleagues for fruitful and enjoyable collaborations, in particular Luigi Delle Site, Rafael Delgado Buscalioni, Davide Donadio, Pep Español, Ralf Everaers, Sebastian Fritsch, Mara Jochum, Christoph Junghans, Biswaroop Mukherjee, Simon Poblete, Matej Praprotnik, and Nico van der Vegt. We would also like to thank the Kavli Institute for Theoretical Physics for sponsoring and hosting the workshop "Physical Principles of Multiscale Modeling, Analysis and Simulation in Soft Condensed Matter" and the participants of this workshop for many stimulating discussions. CP acknowledges financial support by the German Science Foundation within the Emmy Noether Programme (grant PE 1625/1-1) and by the Volkswagen Foundation within the call"New Conceptual Approaches to Modeling and Simulation of Complex Systems".

References

1. R. Potestio, C. Peter, and K. Kremer, "Computer simulations of soft matter: Linking the scales," *Entropy*, vol. 16, no. 8, pp. 4199–4245, 2014.
2. K. Kreis, A. C. Fogarty, K. Kremer, and R. Potestio, "Advantages and challenges in coupling an ideal gas to atomistic models in adaptive resolution simulations," *arXiv:1412.6810 [cond-mat.stat-mech]*, 2014.
3. B. Alder and T. Wainwright, "Phase transition for a hard sphere system," *J. Chem. Phys.*, vol. 5, no. 27, 1957.
4. G. Grest and K. Kremer, "Molecular dynamics simulation for polymers in the presence of a heat bath," *Phys. Rev. A*, vol. 33, pp. 3628–3631, May 1986.
5. K. Kremer, G. Grest, and I. Carmesin, "Crossover from rouse to reptation dynamics: A molecular-dynamics simulation," *Phys. Rev. Lett.*, vol. 61, pp. 566–569, Aug 1988.
6. J. McCammon and M. Karplus, "Internal motions of antibody molecules," *Nature*, vol. 268, pp. 765–766, Aug 1977.
7. M. Karplus and J. McCammon, "Protein structural fluctuations during a period of 100 ps," *Nature*, vol. 277, pp. 578–578, Feb 1979.
8. P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, and M. Parrinello, "Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics," *J Phys Chem B*, vol. 110, pp. 3533–3539, Mar 2006.
9. H. Lou and R. I. Cukier, "Molecular dynamics of apo-adenylate kinase: A distance replica exchange method for the free energy of conformational fluctuations," *J Phys Chem B*, vol. 110, pp. 12796–12808, Jun 2006.
10. K. Arora and C. L. Brooks, "Large-scale allosteric conformational transitions of adenylate kinase appear to involve a population-shift mechanism," *Proc Natl Acad Sci U S A*, vol. 104, pp. 18496–18501, Nov 2007.
11. F. Pontiggia, A. Zen, and C. Micheletti, "Small and large scale conformational changes of adenylate kinase: a molecular dynamics study of the subdomain motion and mechanics," *Biophys J*, vol. 95, pp. 5901–5912, Dec 2008.
12. K. Kremer, "Computer simulations in soft matter science," in *Soft and fragile matter: non equilibrium dynamics, metastability and flow*, vol. 53 of *SUSSP proceedings*, pp. 145–184, IOP Publishing Ltd., 2000.
13. N. A. van der Vegt, C. Peter, and K. Kremer, *Structure-Based Coarse- and Fine-Graining in Soft Matter Simulations*. CRC Press - Taylor and Francis Group, 2009.
14. C. Hijón, E. Vanden-Eijnden, R. Delgado-Buscalioni, and P. Español, "Mori-Zwanzig formalism as a practical computational tool.," *Faraday discussions*, vol. 144, pp. 301–22; discussion 323–45, 467–81, Jan. 2010.
15. W. Noid, *Systematic methods for structurally consistent coarse-grained models*, vol. 924 of *Methods in Molecular Biology*. Humana Press, 2013.
16. W. G. Noid, "Perspective: Coarse-grained models for biomolecular systems," *J. Chem. Phys.*, vol. 139, no. 9, p. 090901, 2013.
17. M. M. Tirion and D. ben Avraham, "Normal mode analysis of g-actin," *JMB*, vol. 230, pp. 186–195, 1993.
18. M. M. Tirion, "Large amplitude elastic motions in proteins from a single–parameter, atomic analysis," *Physical Review Letters*, vol. 77, pp. 1905–1908, 1996.
19. I. Bahar, A. R. Atilgan, and B. Erman, "Direct evaluation of thermal fluctuations in

proteins using a single parameter harmonic potential," *Folding and Design*, vol. 2, pp. 173–181, 1997.
20. C. Micheletti, P. Carloni, and A. Maritan, "Accurate and efficient description of protein vibrational dynamics: comparing molecular dynamics and gaussian models," *Proteins*, vol. 55, pp. 635–645, May 2004.
21. R. Potestio, F. Pontiggia, and C. Micheletti, "Coarse-grained description of proteins' internal dynamics: an optimal strategy for decomposing proteins in rigid subunits," *Biophys J*, vol. 96, 2009.
22. C. Globisch, V. Krishnamani, M. Deserno, and C. Peter, "Optimization of an elastic network augmented coarse grained model to study ccmv capsid deformation," *PLoS ONE*, vol. 8, p. e60582, 04 2013.
23. M. Praprotnik, L. Delle Site, and K. Kremer, "Adaptive resolution molecular-dynamics simulation: Changing the degrees of freedom on the fly," *J. Chem. Phys.*, vol. 123, no. 22, pp. 224106–14, 2005.
24. M. Praprotnik, L. Delle Site, and K. Kremer, "Adaptive resolution scheme for efficient hybrid atomistic-mesoscale molecular dynamics simulations of dense liquids," *Phys. Rev. E*, vol. 73, p. 066701, 2006.
25. M. Praprotnik, L. Delle Site, and K. Kremer, "A macromolecule in a solvent: Adaptive resolution molecular dynamics simulation," *J. Chem. Phys.*, vol. 126, p. 134902, 2007.
26. M. Praprotnik, L. Delle Site, and K. Kremer, "Multiscale simulation of soft matter: From scale bridging to adaptive resolution," *Ann. Rev. Phys. Chem.*, vol. 59, no. 1, pp. 545–571, 2008.
27. S. Fritsch, C. Junghans, and K. Kremer, "Structure formation of toluene around c60: Implementation of the adaptive resolution scheme (adress) into gromacs," *J. Chem. Theory Comput.*, vol. 8, pp. 398–403, FEB 2012.
28. A. B. Poma and L. D. Site, "Classical to Path-Integral adaptive resolution in molecular simulation: Towards a smooth Quantum-Classical coupling," *Phys. Rev. Lett.*, vol. 104, p. 250201, June 2010.
29. R. Potestio and L. Delle Site, "Quantum locality and equilibrium properties in low-temperature parahydrogen: A multiscale simulation study," *J. Chem. Phys.*, vol. 136, FEB 7 2012.
30. B. Ensing, S. Nielsen, P. Moore, M. Klein, and M. Parrinello, "Energy conservation in adaptive hybrid Atomistic/Coarse-Grain molecular dynamics," *J. Chem. Theor. Comp.*, vol. 3, no. 3, pp. 1100–1105, 2007.
31. M. Praprotnik, S. Poblete, L. Delle Site, and K. Kremer, "Comment on "adaptive multiscale molecular dynamics of macromolecular fluids"," *Phys. Rev. Lett.*, vol. 107, p. 099801, Aug 2011.
32. R. Potestio, S. Fritsch, P. Español, R. Delgado-Buscalioni, K. Kremer, R. Everaers, and D. Donadio, "Hamiltonian adaptive resolution simulation for molecular liquids," *Phys. Rev. Lett.*, vol. 110, p. 108301, Mar 2013.
33. R. Potestio, P. Español, R. Delgado-Buscalioni, R. Everaers, K. Kremer, and D. Donadio, "Monte carlo adaptive resolution simulation of multicomponent molecular liquids," *Phys. Rev. Lett.*, vol. 111, p. 060601, Aug 2013.
34. J.-W. Shen, C. Li, N. F. A. van der Vegt, and C. Peter, "Transferability of Coarse Grained Potentials: Implicit Solvent Models for Hydrated Ions," *J. Chem. Theory*

Comput., vol. 7, pp. 1916–1927, 2011.
35. A. Villa, C. Peter, and N. F. A. van der Vegt, "Transferability of Nonbonded Interaction Potentials for Coarse-Grained Simulations: Benzene in Water," *J. Chem. Theory Comput.*, vol. 6, pp. 2434–2444, 2010.
36. B. Mukherjee, L. Delle Site, K. Kremer, and C. Peter, "Derivation of Coarse Grained Models for Multiscale Simulation of Liquid Crystalline Phase Transitions," *J Phys Chem B*, vol. 116, pp. 8474–8484, July 2012.
37. B. Mukherjee, C. Peter, and K. Kremer, "Dual translocation pathways in smectic liquid crystals facilitated by molecular flexibility," *Phys Rev E*, vol. 88, p. 010502, July 2013.
38. C. Lopez, S. Nielsen, P. Moore, J. Shelley, and M. Klein, "Self-assembly of a phospholipid Langmuir monolayer using a coarse-grained molecular dynamics simulations.," *J. Phys.: Condens. Matter*, vol. 14, pp. 9431–9444, 2002.
39. I. R. Cooke, K. Kremer, and M. Deserno, "Tunable generic model for fluid bilayer membranes," *Phys. Rev. E*, vol. 72, no. 1, p. 011506, 2005.
40. M. Müller, K. Katsov, and M. Schick, "Biological and synthetic membranes: What can be learned from a coarse-grained description?," *Physics Reports*, vol. 434, pp. 113 – 176, 2006.
41. B. J. Reynwar, G. Illya, V. A. Harmandaris, M. M. Müller, K. Kremer, and M. Deserno, "Aggregation and vesiculation of membrane proteins by curvature-mediated interactions," *Nature*, vol. 447, no. 7143, pp. 461 – 464, 2007.
42. M. L. Klein and W. Shinoda, "Large-scale molecular dynamics simulations of self-assembling systems," *Science*, vol. 321, pp. 798–800, 2008.
43. N. Go, "Theoretical-studies of protein folding," *Annu. Rev. Biophys. Bioeng.*, vol. 12, pp. 183 – 210, 1983.
44. D. Thirumalai and D. K. Klimov, "Deciphering the timescales and mechanisms of protein folding using minimal off-lattice models," *Curr. Opin. Struct. Biol.*, vol. 9, no. 2, pp. 197 – 207, 1999.
45. A. Liwo, P. Arlukowicz, C. Czaplewski, S. Oldziej, J. Pillardy, and H. A. Scheraga, "A method for optimizing potential-energy functions by a hierarchical design of the potential-energy landscape: Application to the UNRES force field," *Proc. Natl. Acad. Sci. USA*, vol. 99, no. 4, pp. 1937 – 1942, 2002.
46. G. Favrin, A. Irback, and S. Wallin, "Folding of a small helical protein using hydrogen bonds and hydrophobicity forces," *Proteins*, vol. 47, pp. 99–105, 2002.
47. T. Head-Gordon and S. Brown, "Minimalist models for protein folding and design," *Curr. Opin. Struct. Biol.*, vol. 13, no. 2, pp. 160 – 167, 2003.
48. H. D. Nguyen and C. K. Hall, "Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides," *Proc. Natl. Acad. Sci. USA*, vol. 101, no. 46, pp. 16180 – 16185, 2004.
49. N. V. Buchete, J. E. Straub, and D. Thirumalai, "Development of novel statistical potentials for protein fold recognition," *Curr. Opin. Struct. Biol.*, vol. 14, no. 2, pp. 225 – 232, 2004.
50. C. Clementi, "Coarse-grained models of protein folding: toy models or predictive tools?," *Curr. Opin. Struc. Biol.*, vol. 18, pp. 10–15, 2008.
51. P. Derreumaux and N. Mousseau, "Coarse-grained protein molecular dynamics simulations," *J. Chem. Phys.*, vol. 126, p. 025101, 2007.

52. G. Bellesia and J. E. Shea, "Self-assembly of beta-sheet forming peptides into chiral fibrillar aggregates," *J. Chem. Phys.*, vol. 126, p. 245104, 2007.
53. T. Bereau and M. Deserno, "Generic coarse-grained model for protein folding and aggregation," *J. Chem. Phys.*, vol. 130, p. 235106, 2009.
54. V. Tozzini, "Minimalist models for proteins: a comparative analysis," *Q. Rev. Biophys.*, vol. 43, pp. 333–371, 2010.
55. C. Wu and J.-E. Shea, "Coarse-grained models for protein aggregation," *Curr. Opin. Struc. Biol.*, vol. 21, pp. 209–220, 2011.
56. P. de Gennes, *Scaling concepts in polymer physics*. Ithaca [u.a.]: Cornell Univ. Pr., 1979.
57. P. G. De Gennes, "Some conformation problems for long macromolecules," *Reports on Progress in Physics*, vol. 32, no. 2, pp. 187–&, 1969.
58. P. G. De Gennes, "Exponents for excluded volume problem as derived by Wilson method," *Phys. Lett. A*, vol. A 38, no. 5, pp. 339–&, 1972.
59. K. Kremer and G. S. Grest, "Dynamics of entangled linear polymer melts:? a molecular?dynamics simulation," *The Journal of Chemical Physics*, vol. 92, no. 8, pp. 5057–5086, 1990.
60. M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*. Oxford University Press, USA, Nov. 1986.
61. S. O. Nielsen, C. F. Lopez, G. Srinivas, and M. L. Klein, "A coarse grain model for n-alkanes parameterized from surface tension data," *J. Chem. Phys.*, vol. 119, pp. 7043–7049, 2003.
62. S. J. Marrink, A. H. deVries, and A. E. Mark, "Coarse grained model for semiquantitative lipid simulations," *J. Phys. Chem. B*, vol. 108, pp. 750–760, 2004.
63. S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries, "The MARTINI force field: Coarse grained model for biomolecular simulations," *J. Phys. Chem. B*, vol. 111, pp. 7812–7824, 2007.
64. W. Shinoda, R. DeVane, and M. L. Klein, "Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants," *Mol. Simulat.*, vol. 33, pp. 27–36, 2007.
65. L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S. J. Marrink, "The MARTINI coarse-grained force field: Extension to proteins," *J. Chem. Theor. Comput.*, vol. 4, pp. 819–834, 2008.
66. B. M. Mognetti, L. Yelash, P. Virnau, W. Paul, K. Binder, M. Mueller, and L. G. Macdowell, "Efficient prediction of thermodynamic properties of quadrupolar fluids from simulation of a coarse-grained model: The case of carbon dioxide," *J. Chem. Phys.*, vol. 128, p. 104501, 2008.
67. B. M. Mognetti, P. Virnau, L. Yelash, W. Paul, K. Binder, M. Müller, and L. G. Macdowell, "Coarse-grained models for fluids and their mixtures: Comparison of Monte Carlo studies of their phase behavior with perturbation theory and experiment," *J. Chem. Phys.*, vol. 130, p. 044101, 2009.
68. C. A. López, A. J. Rzepiela, A. H. de Vries, L. Dijkhuizen, P. H. Hünenberger, and S. J. Marrink, "Martini Coarse-Grained Force Field: Extension to Carbohydrates," *J. Chem. Theory Comput.*, vol. 5, pp. 3195–3210, 2009.
69. R. DeVane, W. Shinoda, P. B. Moore, and M. L. Klein, "Transferable Coarse Grain Nonbonded Interaction Model for Amino Acids," *J. Chem. Theory Comput.*, vol. 5,

pp. 2115–2124, 2009.
70. R. DeVane, M. L. Klein, C.-C. Chiu, S. O. Nielsen, W. Shinoda, and P. B. Moore, "Coarse-Grained Potential Models for Phenyl-Based Molecules: I. Parametrization Using Experimental Data," *J. Phys. Chem. B*, vol. 114, pp. 6386–6393, 2010.
71. X. He, W. Shinoda, R. DeVane, and M. L. Klein, "Exploring the utility of coarse-grained water models for computational studies of interfacial systems," *Mol. Phys.*, vol. 108, pp. 2007–2020, 2010.
72. S. O. Yesylevskyy, L. V. Schafer, D. Sengupta, and S. J. Marrink, "Polarizable Water Model for the Coarse-Grained MARTINI Force Field," *PLoS Comput. Biol.*, vol. 6, p. e1000810, 2010.
73. W. Tschöp, K. Kremer, J. Batoulis, T. Burger, and O. Hahn, "Simulation of polymer melts. i. coarse-graining procedure for polycarbonates," *Acta Polym.*, vol. 49, no. 2-3, pp. 61 – 74, 1998.
74. A. P. Lyubartsev and A. Laaksonen, "Calculation of effective interaction potentials from radial-distribution functions - a reverse Monte-Carlo approach," *Phys. Rev. E*, vol. 52, pp. 3730 – 3737, 1995.
75. A. P. Lyubartsev and A. Laaksonen, "Osmotic and activity coefficients from effective potentials for hydrated ions," *Phys. Rev. E*, vol. 55, pp. 5689–5696, 1997.
76. F. Müller-Plathe, "Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back," *ChemPhysChem*, vol. 3, pp. 754 – 769, 2002.
77. D. Reith, M. Pütz, and F. Müller-Plathe, "Deriving effective mesoscale potentials from atomistic simulations," *J. Comput. Chem.*, vol. 24, no. 13, pp. 1624–1636, 2003.
78. C. Peter, L. Delle Site, and K. Kremer, "Classical simulations from the atomistic to the mesoscale: coarse graining an azobenzene liquid crystal," *Soft Matter*, vol. 4, pp. 859–869, 2008.
79. T. Murtola, M. Karttunen, and I. Vattulainen, "Systematic coarse graining from structure using internal states: Application to phospholipid/cholesterol bilayer," *J. Chem. Phys.*, vol. 131, p. 055101, 2009.
80. A. Lyubartsev, A. Mirzoev, L. J. Chen, and A. Laaksonen, "Systematic coarse-graining of molecular models by the Newton inversion method," *Faraday Discuss*, vol. 144, pp. 43–56, 2010.
81. A. Savelyev and G. A. Papoian, "Molecular renormalization group coarse-graining of electrolyte solutions: application to aqueous NaCl and KCl," *J. Phys. Chem. B*, vol. 113, pp. 7785–7793, 2009.
82. A. Savelyev and G. A. Papoian, "Molecular Renormalization Group Coarse-Graining of Polymer Chains: Application to Double-Stranded DNA," *Biophys. J.*, vol. 96, pp. 4044–4052, 2009.
83. A. Savelyev and G. A. Papoian, "Chemically accurate coarse graining of double-stranded DNA," *P. Natl. Acad. Sci.*, vol. 107, pp. 20340–20345, 2010.
84. G. Megariotis, A. Vyrkou, A. Leygue, and D. N. Theodorou, "Systematic Coarse Graining of 4-Cyano-4 '-pentylbiphenyl," *Ind. Eng. Chem. Res.*, vol. 50, pp. 546–556, 2011.
85. B. Mukherje, D. S. L., K. K., and C. Peter, "Derivation of a Coarse Grained model for Multiscale Simulation of Liquid Crystalline Phase Transitions," *J. Phys. Chem B submitted*, 2012.
86. S. Izvekov and G. A. Voth, "A multiscale coarse-graining method for biomolecular

systems," *J. Phys. Chem. B*, vol. 109, pp. 2469 – 2473, 2005.
87. G. S. Ayton, W. G. Noid, and G. A. Voth, "Multiscale modeling of biomolecular systems: in serial and in parallel," *Curr. Opin. Struct. Biol.*, vol. 17, pp. 192 – 198, 2007.
88. J. Zhou, I. F. Thorpe, S. Izvekov, and G. A. Voth, "Coarse-grained peptide modeling using a systematic multiscale approach," *Biophys. J.*, vol. 92, pp. 4289 – 4303, 2007.
89. R. D. Hills, L. Lu, and G. A. Voth, "Multiscale Coarse-Graining of the Protein Energy Landscape," *PLoS Comput. Biol.*, vol. 6, p. e1000827, 2010.
90. S. Izvekov, P. W. Chung, and B. M. Rice, "The multiscale coarse-graining method: Assessing its accuracy and introducing density dependent coarse-grain potentials," *J. Chem. Phys.*, vol. 133, p. 064109, 2010.
91. J. W. Mullinax and W. G. Noid, "Recovering physical potentials from a model protein databank," *P. Natl. Acad. Sci. Usa*, vol. 107, pp. 19867–19872, 2010.
92. M. S. Shell, "The relative entropy is fundamental to multiscale and inverse thermodynamic problems," *J. Chem. Phys.*, vol. 129, p. 144108, 2008.
93. A. Chaimovich and M. S. Shell, "Relative entropy as a universal metric for multiscale errors," *Phys. Rev. E*, vol. 81, p. 060104, 2010.
94. A. Chaimovich and M. S. Shell, "Coarse-graining errors and numerical optimization using a relative entropy framework," *J .Chem. Phys.*, vol. 134, p. 094112, 2011.
95. R. Rudd and J. Broughton, "Concurrent coupling of length scales in solid state systems," *Phys. Status Solidi B-Basic Res.*, vol. 217, pp. 251–291, JAN 2000.
96. J. Rottler, S. Barsky, and M. Robbins, "Cracks and crazes: On calculating the macroscopic fracture energy of glassy polymers from molecular simulations," *Phys. Rev. Lett.*, vol. 89, no. 14, p. 148304, 2002.
97. G. Csanyi, T. Albaret, M. C. Payne, and A. D. Vita, ""Learn on the fly": A hybrid classical and Quantum-Mechanical molecular dynamics simulation," *Phys. Rev. Lett.*, vol. 93, no. 17, p. 175503, 2004.
98. D. Jiang and E. Carter, "First principles assessment of ideal fracture energies of materials with mobile impurities: implications for hydrogen embrittlement of metals," *Acta Materialia*, vol. 52, no. 16, pp. 4801–4807, 2004.
99. G. Lu, E. Tadmor, and E. Kaxiras, "From electrons to finite elements: A concurrent multiscale approach for metals," *Phys. Rev. B*, vol. 73, no. 2, p. 024108, 2006.
100. A. Warshel and M. Levitt, "Theoretical studies of enzymic reactions - dielectric, electrostatic and steric stabilization of carbonium-ion in reaction of lysozyme," *J. Mol. Biol.*, vol. 103, pp. 227–249, Jan 1976.
101. J. Gao, K. Lipkowitz, and D. Boyd, *Methods and applications of combined quantum mechanical and molecular mechanical potentials*. Wiley, 1995.
102. M. Svensson, S. Humbel, R. Froese, T. Matsubara, S. Sieber, and K. Morokuma, "Oniom: A multilayered integrated mo+mm method for geometry optimizations and single point energy predictions. a test for diels-alder reactions and pt(p(t-bu)(3))(2)+h-2 oxidative addition," *J. Phys. Chem. Us*, vol. 100, pp. 19357–19363, Jan 1996.
103. P. Carloni, U. Rothlisberger, and M. Parrinello, "The role and perspective of a initio molecular dynamics in the study of biological systems," *Accounts Chem. Res.*, vol. 35, pp. 455–464, Jan 2002.
104. R. Bulo, B. Ensing, J. Sikkema, and L. Visscher, "Toward a practical method for

adaptive qm/mm simulations," *J. Chem. Theor. Comp.*, vol. 5, no. 9, pp. 2212–2221, 2009.
105. L. Delle Site, "Some fundamental problems for an energy-conserving adaptive-resolution molecular dynamics scheme," *Phys. Rev. E*, vol. 76, no. 4, 2007.
106. S. Fritsch, S. Poblete, C. Junghans, G. Ciccotti, L. Delle Site, and K. Kremer, "Adaptive resolution molecular dynamics simulation through coupling to an internal particle reservoir," *Phys. Rev. Lett.*, vol. 108, APR 27 2012.
107. H. Wang, C. Junghans, and K. Kremer, "Comparative atomistic and coarse-grained study of water: What do we lose by coarse-graining?," *Eur. Phys. J. E*, vol. 28, pp. 221–229, 2009.
108. S. Poblete, M. Praprotnik, K. Kremer, and L. Delle Site, "Coupling different levels of resolution in molecular simulations," *J. Chem. Phys.*, vol. 132, no. 11, p. 114101, 2010.
109. D. Mukherji, N. F. A. van der Vegt, and K. Kremer, "Preferential solvation of triglycine in aqueous urea: An open boundary simulation approach," *Journal of Chemical Theory and Computation*, vol. 8, pp. 3536–3541, OCT 2012.
110. L. Delle Site, S. Leon, and K. Kremer, "BPA-PC on a ni(111) surface: The interplay between adsorption energy and conformational entropy for different Chain-End modifications," *J. Am. Chem. Soc.*, vol. 126, pp. 2944–2955, Mar. 2004.
111. B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, "Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation," *J. Chem. Theor. Comput.*, vol. 4, no. 3, pp. 435–447, 2008.
112. J. D. Halverson, T. Brandes, O. Lenz, A. Arnold, S. Bevc, V. Starchenko, K. Kremer, T. Stuehn, and D. Reith, "Espresso++: A modern multiscale simulation package for soft matter systems," *Computer Physics Communications*, vol. 184, no. 4, pp. 1129 – 1149, 2013.
113. B. J. Lambeth, C. Junghans, K. Kremer, C. Clementi, and L. Delle Site, "Communication: On the locality of hydrogen bond networks at hydrophobic interfaces," *The Journal of Chemical Physics*, vol. 133, no. 22, p. 221101, 2010.
114. A. Poma and L. Delle Site, "Adaptive resolution simulation of liquid para-hydrogen: testing the robustness of the quantum-classical adaptive coupling," *Phys. Chem. Chem. Phys.*, vol. 13, pp. 10510–10519, 2011.
115. I. Silvera and V. Goldman, "The isotropic intermolecular potential for h2 and d2 in the solid and gas phases," *J.Chem.Phys.*, pp. 4209–4213, 1978.
116. I. Silvera, "The solid molecular hydrogens in the condensed phase: Fundamentals and static properties," *Rev. Mod. Phys.*, pp. 393–452, 1980.
117. D. Mukherji, N. F. A. van der Vegt, K. Kremer, and L. Delle Site, "Kirkwood-buff analysis of liquid mixtures in an open boundary simulation," *Journal of Chemical Theory and Computation*, vol. 8, no. 2, pp. 375–379, 2012.
118. D. Mukherji and K. Kremer, "Coil-globule-coil transition of pnipam in aqueous methanol: Coupling all-atom simulations to semi-grand canonical coarse-grained reservoir," *Macromolecules*, vol. 46, no. 22, pp. 9158–9163, 2013.
119. H. Wang, C. Hartmann, C. Schütte, and L. Delle Site, "Grand-canonical-like molecular-dynamics simulations by using an adaptive-resolution technique," *Phys. Rev. X*, vol. 3, p. 011018, Mar 2013.
120. A. Heyden and D. G. Truhlar, "Conservative algorithm for an adaptive change of res-

olution in mixed Atomistic/Coarse-Grained multiscale simulations," *J. Chem. Theor. Comp.*, vol. 4, no. 2, pp. 217–221, 2008.
121. J. H. Park and A. Heyden, "Solving the equations of motion for mixed atomistic and coarse-grained systems," *Mol. Sim.*, vol. 35, no. 10–11, pp. 962–973, 2009.
122. M. E. Johnson, T. Head-Gordon, and A. A. Louis, "Representability problems for coarse-grained water potentials," *J. Chem. Phys.*, vol. 126, no. 14, p. 144509, 2007.
123. J. Kirkwood, "Statistical mechanics of fluid mixtures," *J. Chem. Phys.*, vol. 3, no. 5, p. 300, 1935.
124. H. Berendsen, J. Grigera, and T. Straatsma, "The Missing Term in Effective Pair Potentials," *J. Phys. Chem.*, vol. 91, pp. 6269–6271, NOV 19 1987.
125. S. Miyamoto and P. A. Kollman, "Settle: An analytical version of the shake and rattle algorithm for rigid water models," *J. Comput. Chem.*, vol. 13, no. 8, pp. 952–962, 1992.
126. D. Mukherji, N. F. A. van der Vegt, K. Kremer, and L. Delle Site, "Kirkwood-buff analysis of liquid mixtures in an open boundary simulation," *J. Chem. Theory Comput.*, vol. 8, no. 2, pp. 375–379, 2012.

MD of Biomolecules Steered with Mean Solvation Forces Obtained from 3D-RISM-KH Molecular Theory of Solvation

Andriy Kovalenko[1,2] **and Ihor Omelyan**[1,2,3]

[1] National Institute for Nanotechnology, National Research Council of Canada
11421 Saskatchewan Dr., Edmonton, Alberta, T6G 2M9, Canada

[2] Department of Mechanical Engineering, University of Alberta
Mechanical Engineering Bldg. 4-9, Edmonton, Alberta, T6G 2G8, Canada

[3] Institute for Condensed Matter Physics, National Academy of Sciences of Ukraine
1 Svientsitskii Street, Lviv 79011, Ukraine

E-mail: andriy.kovalenko@ualberta.ca

Statistical mechanics based, integral equation theory of molecular liquids is promising as an essential part of multiscale methodology for chemical and biomolecular nanosystems in solution. Beginning with a molecular interaction potential force field, it uses diagrammatic analysis of the solvation free energy to derive integral equations for correlation functions between molecules in solution in the statistical-mechanical ensemble. The infinite chain of coupled integral equations for many-body correlation functions is reduced to a tractable form for 2- or 3-body correlations by applying so-called closure relations. The solvation structure obtained by solving them is of accuracy comparable to molecular simulations that *have converged* but has a critical advantage of readily treating effects and processes spanning over large space and slow time scales, by far not feasible for explicit solvent molecular simulations. One of the versions of this formalism, the three-dimensional reference interaction site model (3D-RISM) integral equation complemented with the Kovalenko-Hirata (KH) closure approximation, yields the solvation structure in terms of 3D maps of density distributions of solvent interaction sites around a solute molecule with full consistent account for effects of chemical functionalities of all solution species. The solvation free energy and subsequent thermodynamics is then obtained at once as a simple integral of the correlation functions by performing thermodynamic integration analytically. The latter allows analytical differentiation of the free energy functional and thus self-consistent field coupling of the 3D-RISM-KH molecular theory of solvation in multiscale approaches. One of such hybrids is multitime step molecular dynamics (MTS-MD) of biomolecules steered with 3D-RISM-KH mean solvation forces. To speed up such hybrid simulation, mean solvation forces acting on atoms of the biomolecule in conformations at successive inner time steps are calculated by the new algorithm of generalized solvation force extrapolation (GSFE) using a relatively small number of best (closest) solute atomic coordinates and corresponding mean solvation forces obtained at previous outer time steps by converging the 3D-RISM-KH integral equations. GSFE is based on a set of techniques including the non-Eckart like transformation of coordinate space separately for each solute atom, extension of the force-coordinate pair basis set followed by selection of the best subset, balancing the normal equations to minimize the deviations by modified least squares, and incremental increase of the outer time step in motion integration. The MTS-MD evolution steered with GSFE of 3D-RISM-KH mean solvation forces is efficiently stabilized with the new method of optimized isokinetic Nosé-Hoover chain (OIN) thermostat. The hybrid OIN/GSFE/3D-RISM-KH integrator has been validated on solvated chemical molecules and biomolecules of different stiffness and complexity: asphaltene dimer in toluene solvent, and hydrated alanine dipeptide, miniprotein 1L2Y, and protein G. The accuracy of GSFE and the efficiency of OIN allows huge outer time step up to 1–4 ps while accurately reproducing conformational properties. Quasidynamics steered with 3D-RISM-KH mean solvation forces results in time scale compression of protein conformational changes and thus in further significant acceleration of sampling with respect to real time dynamics. Overall, a 50– to 1000–fold effective speedup of sampling is achieved, compared to conventional MD with explicit solvent. This enables, for example, to quickly fold the miniprotein from a fully denatured, extended state in 60ns quasidynamics steered with 3D-RISM-KH mean solvation forces, compared to the average physical folding time of 4–9 μs observed in experiment.

0 Prior to Introduction: *Quo vadis*

Nanoscale properties, phenomena, and processes are profoundly different from macroscopic laws governing the behaviour of continuous media and materials. All functional features of nanostructures stem from the microscopic properties of constituting atoms but manifest at length scale from one to hundreds nanometers and time scale up to microseconds and more. By changing size, composition, and fabrication protocol, nanostructure properties and processes can be tuned in a widest range.[1,2] Therefore, predictive modeling of nanosystems must operate at length scales from an Angström to hundreds nanometers and microns and time scales to miliseconds and seconds, and yet derive their properties from the chemical functionalities of the constituents. (An example is biological cellular systems of protein nanomachines operating in crowded environment.) Explicit molecular modeling of such nanosystems involves millions and billions of molecules and is by far not feasible in a "brute force" approach employing just *ab initio* quantum chemical methods and/or molecular simulations. Addressing this challenge requires development and use of multiscale methods coupling several levels of description, from electronic structure methods for building blocks and classical molecular simulations for critical aggregates in the system, to statistical-mechanical theories for their large assemblies and mean properties in a statistical ensemble over characteristic size and time scales, to eventually come up with macroscopic scale properties of the nanostructures and related processes showing up in the "real observable world". A true, genuine challenge of multiscale modeling is development of a theoretical framework that couples methods at different scales, so that observables at lower-level scales are **analytically** linked in a self-consistent field description to force fields of more coarse-grained models at higher-level scales.[3] Statistical mechanics itself, in particular, integral equation theory of liquids,[4] is an example of such a theoretical coupling between microscopic molecular variables and thermodynamic, macroscopic properties.

1 Motivation and Introduction

Prediction of the structure and functioning of proteins at the molecular level from amino acids sequences and interaction potentials between the solute and solvent atoms remains a challenging task.[5] The main problem is that conformational and folding equilibria in these complex systems take place on time scales ranging from picoseconds to micro- and milliseconds. Molecular dynamics (MD) simulations originally introduced to study simple liquids[6–9] have been developed into a powerful tool[10–22] that enabled understanding of the mechanisms of protein folding, one of the most fundamental biochemical operations. However, MD must at least several microseconds long to stand a good chance of observing a single folding event even for the simplest proteins.[5] In the course of development, MD simulation length for relatively simple systems reached one microsecond,[12,13] ten microseconds,[18] and recently one millisecond.[19–21] However, MD simulations of relatively large proteins are restricted typically to tens to hundreds of nanoseconds,[23] which is quite insufficient for usual MD to obtain a full pattern on the folding behavior.

A promising way to significantly accelerate molecular simulations has been to combine the MD method with the 3D-RISM integral equation theory of molecular liquids (three-dimensional reference interaction site model)[24–33] complemented with the Kovalenko-

Hirata closure relation.[28,31,33] In the hybrid MD/3D-RISM-KH approach, individual trajectories and dynamics of solvent molecules are contracted to quasiequilibrium 3D density distribution functions of their interaction sites around the biomolecule in successive conformation snapshots. The evolution of the biomolecule thus becomes quasidynamics steered with mean solvation forces obtained for each conformation of the biomolecule from the 3D-RISM-KH molecular theory of solvation.[34–38] The latter produces mean solvation forces by converging the 3D-RISM-KH integral equations derived from the first principles of statistical mechanics, beginning from an input of the interaction potentials and geometries of the biomolecule and solvent molecules (molecular force field). The 3D-RISM-KH mean solvation forces statistically-mechanically averaged over the distributions of an infinite number of solvent molecules are thus added to the direct intramolecular interactions for integrating the equations of motion of atoms in the biomolecule. A chief advantage of this hybrid approach is that slow solvation processes in the confined geometry of the biomolecule, particularly under its conformational changes, such as solvent exchange and re-equilibration, localization of structural solvent, ions distribution and localization, and protein-ligand binding, which constitute a major challenge for conventional MD are readily accounted for by 3D-RISM-KH mean solvation forces and thus excluded from the biomolecule quasidynamics. This leads to drastic compression of time scale in protein quasidynamics compared to real dynamics, and thus enables fast access to structural and folding properties of large biomolecular systems in solution.

The 3D-RISM-KH molecular theory of solvation has a profound advantage compared to continuum solvation methods which represent polar solvation forces with either the Poisson-Boltzmann (PB)[39] or the Generalized Born (GB)[40–42] models and empirically account for nonpolar solvation forces with the solvent accessible surface area (SASA, or SA) model supplemented with additional volume and dispersion integral terms.[43,44] These continuum solvation approaches are parameterized for hydration of biomolecules and are not transferable to other solvent or solvent system with cosolvent, ions at a finite concentration (physiological concentration in biomolecular systems), and other solvent species, in particular, in the recent methods[45–48] treating ligand fragments as part of a solvent mixture. Continuum solvation models based on the concept of a solvation cavity in dielectric structureless medium representing solvent entirely ignore effects of finite size of solvent molecules on mean solvation forces between solute molecules, for example, a desolvation barrier due to expelling solvent molecules from the gap between the surfaces of proteins (parts of protein) when bringing them together in contact. Furthermore, SASA is well defined for an outer surface of a biomolecule but loses physical meaning and becomes inadequate inside small inner cavities of biomolecules like a narrow channel accommodating an ion and a few water molecules.

As distinct, the 3D-RISM-KH molecular theory of solvation readily accounts for molecular solvation effects and yields the solvation structure and thermodynamics at the level of molecular simulations. (A comparison attainable if molecular simulations are feasible; for many biomolecular systems, affordable simulation times are too short to gain meaningful statistics of slow essential solvation events, while the molecular theory of solvation provides the solvation structure and thermodynamics in the equilibrium ensemble.) In the context of solvation models nomenclature used in the literature, the GB(PB)SA models are both *implicit* and *continuum*, whereas the 3D-RISM-KH theory is *implicit* as it produces 3D density distributions which are an average of individual trajectories of solvent

molecules but is *not continuum*. With an input of all-atom interaction potentials (molecular force field), the 3D-RISM-KH theory produces the full molecular picture of solvation at output in terms of 3D maps of solvent density distributions and solvation free energy density, including energy-entropy decomposition, with full account of chemical specificities and molecular geometries specified in the force field.

A further, even more important advantage of the hybrid MTS-MD/3D-RISM-KH approach is that the slow components of motion due to solvent exchange in confined spaces and re-equilibration in conformational changes of the solute biomolecule are entirely eliminated from the MTS-MD/3D-RISM-KH quasidynamics. This drastically contracts the quasidynamics time scale compared to real dynamics, and dramatically shorter simulations are required to gain adequate statistics. Moreover, this intrinsic acceleration of MD steered with 3D-RISM-KH mean solvation forces grows with complexity of the biomolecule (e.g., protein). Indeed, in conventional MD, solvent enters pockets and inner cavities of the biomolecule through its conformational changes. This is a very slow process with rare statistics which is as difficult to model explicitly as protein folding conformational changes themselves. (Note again that continuum solvation models do not adequately reproduce solvation in inner cavities of biomolecules.) As distinct, the 3D-RISM-KH theory yields the solvent distribution in the inner cavity or pocket at once for the final conformation in chemical equilibrium with the bulk solvent outside the protein by construct of the theory, bypassing the intermediate conformational states.

Calculation of the 3D-RISM-KH mean solvation forces requires significant numerical efforts if the integral equations are solved at each inner time step of the MD trajectory. However, 3D-RISM-KH mean solvation forces vary with solute atomic coordinates and so with time much smoother than solute-solvent atomic interaction forces evaluated directly in conventional MD. The reason is that mean solvation forces at a given solute conformation are obtained by statistical averaging over all arrangements of equilibrated solvent molecules, and so all core repulsion forces and other strong short-range components typical to explicit solvent interactions are smoothed out in the averaging. Therefore, 3D-RISM-KH forces can be efficiently extrapolated, which allows the 3D-RISM-KH integral equations to be converged much less frequently and thus drastically increases the efficiency of hybrid MD/3D-RISM-KH simulation.

Miyata and Hirata performed a pioneering MD/3D-RISM-KH simulation for hydrated acetylacetone.[34] They exploited the standard reference system propagator algorithm (RESPA)[49–51] in the microcanonical ensemble, with calculation of mean solvation forces by converging the 3D-RISM-KH integral equation at each outer step without resorting to extrapolation of solvation forces. With this propagator algorithm, it was impossible to apply outer time steps larger than 5 fs as a result of resonance instabilities[52–57] that appear in MD/3D-RISM-KH as well as in conventional MD simulations due to the multiple time step (MTS) interplay between strong intramolecular (solute-solute) and weak intermolecular (solute-solvent) forces. In conventional MD, the accuracy of MTS simulation can be increased by carrying out processed phase-space transformations.[58,59] Utilizing them within an energy-constrained scheme, it was shown[60] in MD simulations of water that outer time steps up to 16 fs are acceptable. However, such steps cannot exceed the theoretical limiting value of 20 fs inherent in the microcanonical description. Furthermore, in hybrid MD/3D-RISM-KH without solvation force extrapolation, the procedure of converging the 3D-RISM-KH integral equations has to be repeated too frequently (every 5 fs),[34] which

drastically slows down the computations.

In order to damp MTS instabilities, the MD/3D-RISM-KH approach has been extended[35] to the canonical ensemble within the Langevin dynamics.[61,62] Introducing the method of solvation force extrapolation (SFE) for mean solvation forces acting on the solute, it has been demonstrated for hydrated alanine dipeptide that larger outer time steps of up to 20 fs are feasible.[35] They, however, are still smaller than those achievable in conventional MD simulations by the best previously known isokinetic Nosé-Hoover chain RESPA (INR) integrator, for which an outer time step of 100 fs is possible.[63,64]

MTS instabilities in MD simulations eliminated much more efficiently by using the optimized isokinetic Nosé-Hoover chain (OIN) canonical ensemble introduced recently.[36] OIN improves over the INR method[63,64] and other canonical-isokinetic schemes[65–67] by coupling each set of Nosé-Hoover chain thermostats to some optimal number of degrees of freedom in the system. With some modification to the original SFE scheme,[35] the OIN integrator has been combined with the MD/3D-RISM-KH approach. It was illustrated on hydrated alanine dipeptide[36] that the OIN ensemble is superior to the Langevin and INR schemes, as it enables large outer time steps of order of several hundred femtoseconds and thus provides a speedup up to 20 times compared to conventional MD with explicit solvent.

The next improvement constituted a method of advanced solvation force extrapolation (ASFE) of mean solvation forces in hybrid MD/3D-RISM-KH simulations, with the pivoting idea to apply a global non-Eckart-like rotation of atomic coordinates to minimize the distances between the biomolecule sites in different conformations at successive time steps.[37] Carrying out the force extrapolation in the transformed space and extending the set of outer conformations evaluate mean solvation forces with a much better accuracy than the previous extrapolation scheme. This allows huge outer time steps up to tens of picoseconds in hybrid MD/3D-RISM-KH simulations without affecting the equilibrium and conformational properties, and resulted in a 100- to 500-fold acceleration compared to MD with explicit solvent for a relatively small system of hydrated alanine dipeptide.[37] The ASFE ideas and techniques have been generalized to larger solute molecules, including proteins, in a method of generalized solvation force extrapolation (GSFE).[38] First of all, a full rotational transformation of the whole molecule has been replaced with an individual non-Eckart-like transformation for each atom of the biomolecule in the presence of a smooth weighing function. For macromolecules, this appreciably enhances convergence of the extrapolated forces to their exact values with increasing the number of basis outer coordinates. Modifications have included also the numerical techniques of least-squares minimization, extension of the force-coordinate pair set to select the best subset, and balancing the normal equations. In addition, the so-called frequency scheme has been proposed, which appreciably reduces computational overhead of the extrapolation without loss of precision. With the above improvements, the good accuracy attainable with the original SFE scheme[35] for outer time steps up to 20 fs can now be held with the GSFE method for much longer steps of order of 1 to 4 ps, even for large biomolecules.

2 MD of a Biomolecule Steered with Mean Solvation Forces

Let's consider a solute macromolecule consisting of M atoms solvated in liquid comprising a large number of solvent molecules with M' atomic sites. Conventional MD simulation deals with forces $-\partial U/\partial \mathbf{r}_i$ acting on all atoms $i = 1, \ldots, M + M'$ at positions $\mathbf{r}_i$ in the

solute-solvent system with the total potential energy $U(\mathbf{r}_1, \ldots, \mathbf{r}_M, \mathbf{r}_{M+1}, \ldots, \mathbf{r}_{M+M'})$. The latter can be split up as $U = U_1 + U_2$ into the solute-solute interaction potential $U_1(\mathbf{r}_1, \ldots, \mathbf{r}_M)$ and the remaining term $U_2(\mathbf{r}_1, \ldots, \mathbf{r}_{M+M'})$ comprising the solute-solvent and solvent-solvent interaction potentials. Accordingly, the forces acting on solute atoms split up into two parts coming respectively from solute-solute and solute-solvent atomic interactions, $-\partial U/\partial \mathbf{r}_i = -\partial U_1/\partial \mathbf{r}_i - \partial U_2/\partial \mathbf{r}_i$, where $i = 1, \ldots, M$. In MD simulations of biomolecules, the total number of atoms of solvent molecules has to be much larger than that of the solute biomolecule ($M' \gg M$) to have good statistics and neglect finite size effects. Even at infinite dilution, for a biomolecule of $M \sim 1,000 - 10,000$ atoms the number of solvent atoms has to be $M \sim 10,000 - 100,000$. This considerably complicates conventional MD because a large portion of the computational costs is spent on evaluation of solute-solvent and solvent-solvent atomic forces $-\partial U_2/\partial \mathbf{r}_i$, where $i = 1, \ldots, M + M'$. In common practice, the concentration of solute macromolecules (or aggregates of macromolecules) is small and interactions between solutes (or composite solutes) are neglected, thus reducing the consideration to infinite dilution.

Since we are interested exclusively in conformational and folding behavior of the solute biomolecule, one way to improve the efficiency of MD simulations consists in contracting the degrees of freedom of solvent (a huge number of molecules) and evaluating the dynamics of the biomolecule on the solvation free energy surface. Instead of solute-solvent forces $-\partial U_2/\partial \mathbf{r}_i$, $i = 1, \ldots, M$, the dynamics of the biomolecule is then steered with the mean solvation forces[68,69] that are defined as a statistical average of solute-solvent atomic forces $-\partial U_2/\partial \mathbf{r}_i$ acting on each solute atom $i = 1, \ldots, M$ over all arrangements of all M' solvent atoms around the biomolecule at a frozen conformation $\{\mathbf{r}_1, \ldots, \mathbf{r}_M\}$,

$$\mathbf{f}_i(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_M) = -\frac{\int d\mathbf{r}_{M+1} d\mathbf{r}_{M+2} \ldots d\mathbf{r}_{M+M'} \, \frac{\partial U_2}{\partial \mathbf{r}_i} \, e^{-\frac{U}{k_\mathrm{B}T}}}{\int d\mathbf{r}_{M+1} d\mathbf{r}_{M+2} \ldots d\mathbf{r}_{M+M'} \, e^{-\frac{U}{k_\mathrm{B}T}}}, \qquad (1)$$

where k_B is the Boltzmann constant and T is the system temperature. Without explicit solvent treatment, mean solvation forces can be obtained either from continuum solvation models or from molecular theory of solvation.

The 3D-RISM-KH molecular theory of solvation[28–33] is transferable and yields solvation structure and thermodynamics, including mean solvation forces, from the first principles of statistical mechanics with an accuracy (up to the closure approximation used) at the level of explicit solvent molecular simulation **that has converged**. (Viable statistical sampling in explicit solvent simulation is a huge challenge for systems with slow dynamics of solvent exchange in confined spaces, e.g. protein pockets and cavities, structural solvent localization, preferential adsorption of co-solvent, partitioning of ions, binding of ligands, and solvent and ions mediated protein-protein interactions.) On the other hand, as discussed in the Introduction, continuum solvation models suffer from inherent disadvantages: non-transferable to other solvents and solvent systems, in particular, electrolyte solutions; missing solvent size effects such as desolvation barrier in protein aggregation; inadequate to reproduce solvation of internal cavities of biomolecules. Mention again another important advantage that quasidynamics steered with 3D-RISM-KH mean solvation forces achieves time scale compression of protein conformational changes coupled with solvent exchange, which results in significant acceleration of protein conformational sam-

pling with respect to real time dynamics. The 3D-RISM-KH theory has been validated on prediction of the solvation structure and thermodynamics of various chemical species and biomolecules in different solvents,[31,33,70–73] in particular, in the context of hybrid MD/3D-RISM-KH simulations for hydrated biomolecules.[34–38] The 3D-RISM-KH mean solvation forces $\mathbf{f}_i$ can then be employed together with the direct intramolecular solute-solute interactions $-\partial U_1/\partial \mathbf{r}_i$ to integrate the equations of motion just for solute atoms.

3 3D-RISM-KH Molecular Theory of Solvation

The 3D-RISM-KH theory[28–33] yields the solvation structure in terms of the normalized 3D density distribution functions $g_\alpha(\mathbf{r})$ of interaction site α of solvent molecules at spatial position $\mathbf{r}$ around the whole solute macromolecule or supramolecule, starting from the input of intermolecular potentials for explicit solvent and solute molecules (molecular force field). The 3D-RISM integral equation for 3D solute-solvent site correlation functions[24–28,31] can be derived either within the site formalism of density functional theory of molecular liquids[24–26] or from the six-dimensional Ornstein-Zernike integral equation for molecular liquids[4] by orientations averaging centered at interaction sites of solvent molecules to contract orientational degrees of freedom of solvent.[27,28,31] It reads

$$h_\alpha(\mathbf{r}) = \sum_\gamma \int d\mathbf{r}' c_\gamma(\mathbf{r}-\mathbf{r}')\chi_{\gamma\alpha}(r')\,, \tag{2}$$

where $h_\alpha(\mathbf{r})$ and $c_\alpha(\mathbf{r})$ are respectively the 3D total and direct correlation functions of solvent site α around the solute molecule, $\chi_{\gamma\alpha}(r)$ is the radially dependent site-site susceptibility function of solvent which is an input to 3D-RISM and is calculated beforehand, and the indices γ and α enumerate all interaction sites on all sorts of solvent species. Diagrammatic analysis of the total and direct correlation functions[4] relates the former to the density distribution function as $h_\alpha(\mathbf{r}) = g_\alpha(\mathbf{r}) - 1$, and so $h_\alpha(\mathbf{r})$ has the meaning of a normalized 3D distribution of spatial correlations, or normalized deviations of solvent site density around the solute molecule from the average value in the solution bulk. The long-range asymptote of the 3D direct correlation function $c_\alpha(\mathbf{r})$ is given by the 3D interaction potential $u_\alpha(\mathbf{r})$ scaled by $k_\mathrm{B}T$ between the whole solute molecule and solvent interaction site α: $c_\alpha(\mathbf{r}) \sim -u_\alpha(\mathbf{r})/(k_\mathrm{B}T)$ for $\mathbf{r}$ outside the short-range repulsive core region (typically comprising the repulsive slope down to the attractive well minimum); the values of $c_\alpha(\mathbf{r})$ inside the repulsive core are related to the solvation free energy. Usually but not necessarily, the 3D solute-solvent site interaction potential is given by a sum of pairwise isotropic potentials (typically Coulomb and Lennard-Jones) between solute and solvent interaction sites, $u_\alpha(\mathbf{r}) = \sum_i u_{i\alpha}(|\mathbf{r}-\mathbf{r}_i|)$, where $\mathbf{r}_i$ is the location of solute atom i and $\mathbf{r}$ is that of solvent site α.

The 3D-RISM integral equation eq. (2) involves two correlation functions, $h_\alpha(\mathbf{r})$ and $c_\alpha(\mathbf{r})$, and to be complete has to be complemented with another relation between $h_\alpha(\mathbf{r})$ and $c_\alpha(\mathbf{r})$ called a closure which also involves the interaction potential $u_\alpha(\mathbf{r}_\alpha)$ specified at input with the molecular force field. The exact closure relation has a nonlocal functional form that can be represented as an infinite diagrammatic series in terms of multiple integrals of the total correlation function;[4] however, it is computationally intractable, as the series is poorly convergent and the higher-order diagrams are integrals extremely cumbersome to calculate. Therefore, the exact closure is replaced in practice with amenable

approximations which should analytically ensure asymptotics of the correlation functions and features of the solvation structure and thermodynamics to properly represent the solvation physics. The approximation proposed by Kovalenko and Hirata (KH closure),[28,31,33] the 3D version of which reads as

$$g_\alpha(\mathbf{r}) = \begin{cases} \exp\bigl(-\frac{1}{k_\mathrm{B}T}u_\alpha(\mathbf{r}) + h_\alpha(\mathbf{r}) - c_\alpha(\mathbf{r})\bigr) & \text{for } g_\alpha(\mathbf{r}) \leqslant 1 \\ 1 - \frac{1}{k_\mathrm{B}T}u_\alpha(\mathbf{r}) + h_\alpha(\mathbf{r}) - c_\alpha(\mathbf{r}) & \text{for } g_\alpha(\mathbf{r}) > 1 \end{cases} . \tag{3}$$

couples in a nontrivial way the so-called hypernetted chain (HNC) and mean spherical approximation (MSA) closures,[4] the former automatically applied to spatial regions of density depletion $g_\alpha(\mathbf{r}) < 1$, including the repulsive core, and the latter to spatial regions of solvent density enrichment $g_\alpha(\mathbf{r}) > 1$, such as association peaks and long-range tails of near-critical fluid phases, while keeping the right long-range asymptotics of $c_\alpha(\mathbf{r})$ peculiar in both the HNC and MSA. (The distribution function and its first derivative are continuous at the joint boundary $g_\alpha(\mathbf{r}) = 1$ by construct.) The KH approximation consistently accounts for both electrostatic and non-polar solvation forces, such as hydrogen bonding and other associative effects, hydrophobic hydration and interaction, preferential solvation, desolvation and other steric effects for macromolecules and supramolecules in simple and complex liquids, solvent mixtures, non-electrolyte and electrolyte solutions in various chemical,[28–31,33,71–78] soft matter,[79] synthetic organic supramolecular,[33,80–82] biopolymeric,[83–85] and biomolecular[31,33,45–48,70,79,86–96] systems.

The radially dependent site-site susceptibility of solvent $\chi_{\gamma\alpha}(r)$ determines nonlocal response of solvent to an external field, statistically-mechanically averaged over orientations and arrangements of solvent molecules in the solvation shells. In the context of eq. (2), the solvent density change given by $h_\alpha(\mathbf{r})$ comes from the insertion of a solute molecule characterized with $c_\alpha(\mathbf{r})$ which propagates across the solvation shells through the effective solvent-solvent correlations given by the solvent susceptibility $\chi_{\gamma\alpha}(r)$. The latter breaks up into the intra- and intermolecular terms,

$$\chi_{\gamma\alpha}(r) = \omega_{\gamma\alpha}(r) + \rho_\gamma h_{\gamma\alpha}(r), \tag{4}$$

where the intramolecular correlation function $\omega_{\gamma\alpha}(r)$ normalized as $\int d\mathbf{r}\, \omega_{\gamma\alpha}(r) = 1$ represents the geometry of solvent molecules (i.e., $\omega_{\gamma\alpha}(r) = 0$ for sites γ and α on different species). For rigid molecular species with site separations $l_{\gamma\alpha}$ it has the form $\omega_{\gamma\alpha}(r) = \delta(r - l_{\gamma\alpha})/(4\pi l_{\gamma\alpha}^2)$ specified in the reciprocal k-space as $\omega_{\gamma\alpha}(r) = j_0(kl_{\gamma\alpha})$, where $j_0(x)$ is the zeroth-order spherical Bessel function. The radial site-site total correlation function $h_{\gamma\alpha}(r)$ for all pairs of sites on all species of the solvent are obtained in advance to the 3D-RISM-KH calculations from the dielectrically consistent RISM theory[97,98] coupled with the KH closure relation for the radial correlation functions (DRISM-KH approach).[31–33] The DRISM-KH theory can be applied to solution systems of a given composition in a wide range of thermodynamic conditions, including different solvents,[99,100] solvent mixtures,[101,102] polymeric solutions,[79,103] and electrolyte solutions.[31,33,104]

An important feature of the KH closure eq. (3) is that the solvation free energy μ_solv as determined by Kirkwood's thermodynamic integration gradually switching the interactions on from 0 to the full potential $u_\alpha(\mathbf{r})$ is obtained analytically in a closed form of a single spatial integral in terms of the correlation functions:[28,31,33]

$$\mu_{\mathrm{solv}} = \sum_{\alpha} \int_{V} d\mathbf{r} \Phi_{\alpha}(\mathbf{r}), \tag{5a}$$

$$\Phi_{\alpha}(\mathbf{r}) = \rho_{\alpha} k_{\mathrm{B}} T \left[\frac{1}{2} \big(h_{\alpha}(\mathbf{r}) \big)^2 \Theta\big(- h_{\alpha}(\mathbf{r}) \big) - \frac{1}{2} h_{\alpha}(\mathbf{r}) c_{\alpha}(\mathbf{r}) - c_{\alpha}(\mathbf{r}) \right], \tag{5b}$$

where the sum goes over all the sites of all solvent species, and $\Theta(x)$ is the Heaviside step function. This offers several substantial advantages over molecular simulations requiring free energy perturbation techniques with multiple productive runs along the thermodynamic integration or solute mutation which are extremely time consuming.[105] Once the 3D-RISM-KH equations eqs. (2) and (3) are converged for the correlation functions $c_{\alpha}(\mathbf{r})$ and $h_{\alpha}(\mathbf{r})$, all the solvation thermodynamics is readily available from the expression eq. (5) and analysis of its components. The solvation free energy in the form eq. (5) can be split up into partial contributions from each solvent species by grouping the terms for the corresponding sites. It is also possible to resolve partial contributions of different functional groups of the solute macromolecule or supramolecule to the solvation free energy, or so-called spatial decomposition analysis.[74,94] Further, partial molar volume of biomolecules is obtained analytically within the Kirkwood-Buff theory as a simple integral in terms of $c_{\alpha}(\mathbf{r})$, well reproducing experimental data and pressure effects on conformational transitions of proteins (see literature cited in ref.[96]). Other thermodynamic quantities can also be obtained analytically by taking the corresponding derivaties. In particular, mean solvation forces acting on each atom of the solute macromolecule or supramolecule are readily obtained as a simple 3D spatial integral in terms of the 3D site distribution functions $g_{\alpha}(\mathbf{r})$ (see the next section).

The function $\Phi_{\alpha}(\mathbf{r})$ given by eq. (5b) is interpreted as a 3D map of solvation free energy density (3D-SFED) coming from interaction site α of solvent around the solute molecule. The solvation free energy μ_{solv} results from summation of 3D-SFED partial contributions of all solvent species and spatial integration over the whole space V comprising the solvation shells, as well as the solute-solvent molecular repulsive cores. (Inside the cores, the correlation functions are nonzero and related to the free energy of solvent cavity formation to accommodate the solute molecule.) While 3D maps of site density distributions $g_{\alpha}(\mathbf{r})$ provide solvation structure and characterize the strength of attachment of solvent species at the solute molecule surface in terms of potentials of mean force,[33,93] such information on solute-solvent effective interactions is convolved. As distinct, 3D-SFED maps $\Phi_{\alpha}(\mathbf{r})$ spatially resolve the intensity of mean solvation forces statistically-mechanically averaged over solvent arrangements and orientations around the solute molecule and indicate where they contribute the most or the least to the entire solvation free energy.[78,83,85]

To properly treat electrostatic forces in electrolyte solution with polar molecular solvent and ionic species, the long-range electrostatic asymptotics of both the 3D direct and total correlation functions in the 3D-RISM integral equation eq. (2) are separated out and handled analytically.[29–32,73,89,106] The remaining short-range parts of the 3D site correlation functions are discretized on a uniform 3D rectangular grid in a box large enough to accommodate the solvation structure, typically 2 to 3 solvation shell oscillations. The spatial convolution of the short-range therm in eq. (2) is calculated by means of 3D fast Fourier transform. Note that even though the solvent susceptibility $\chi_{\gamma\alpha}(r)$ has a long-range electrostatic part, no aliasing occurs in the backward 3D-FFT of the short-range part of $h_{\alpha}(\mathbf{k})$ on the 3D box supercell since, for the physical reason, it typically contains merely

2-3 oscillations and thus vanishes at the box boundaries.[32] The same analytical treatment with separation of the electrostatic asymptotes is applied also to the radial site-site correlation functions in the DRISM-KH integral equations that produce the water susceptibility eq. (4), as well as to the 3D site correlation functions in the solvation free energy integral eq. (5) which is reduced to a 3D integral of the short-range terms on the 3D box and one-dimensional integrals of the asymptotics easy to compute.[29–32,73,89] Further, the core-shell-asymptotics treatment of solvation shells in MDIIS decreases memory and corresponding CPU load up to an order of magnitude.[32]

The 3D-RISM-KH integral equations eqs. (2) and (3) are converged by using the modified algorithm of direct inversion in the iterative subspace (MDIIS).[107,108,29–32] The MDIIS numerical solver accelerates convergence of integral equations of liquid state theory by optimizing each iterative solution in a Krylov subspace of typically last 10-20 successive iterations and then making the next iterative guess by mixing the optimized solution with the approximated optimized residual. The computational expenses of converging the 3D-RISM-KH equations can be further reduced with several strategies, including a high-quality initial guess for the 3D direct correlation functions $c_\alpha(\mathbf{r})$; pre- and post-processing of the 3D solute-solvent potentials $u_\alpha(\mathbf{r})$, the long-range asymptotics of the 3D correlation functions $c_\alpha(\mathbf{r})$, and $h_\alpha(\mathbf{r})$, and forces; several cutoff schemes and an adaptive solvation box;[35] and the core-shell-asymptotics treatment of solvation shells.[32]

4 Combining MD with 3D-RISM-KH

Unlike conventional MD dealing with trajectories of explicit solvent molecules, the hybrid MD/3D-RISM-KH approach[34–36] contracts them to 3D site density distribution functions $g_\alpha(\mathbf{r})$ of quasiequilibrium solvent at successive conformations of the biomolecule and thus performs quasidynamics of the biomolecule steered with mean solvation forces. The latter can be determined in a general case from Kirkwood's thermodynamic charging integral by differentiation with respect to solute atomic coordinates. For the solvation free energy not dependent on a thermodynamic integration path (which is true for the exact solvation free energy but not necessarily for a given closure approximation of integral equation theory of liquids), the mean solvation force is immediately obtained as the "detailed" solute-solvent site interaction potential force averaged over the solvation shells with the 3D solute-solvent site density distribution function,[34,35]

$$\mathbf{f}_i \equiv \mathbf{f}(\mathbf{r}_i) = -\frac{\partial \mu_{\mathrm{solv}}}{\partial \mathbf{r}_i} = \sum_\alpha \rho_\alpha \int d\mathbf{r}\, g_\alpha(\mathbf{r}) \frac{\partial u_{i\alpha}(\mathbf{r}-\mathbf{r}_i)}{\partial \mathbf{r}_i}\,, \tag{6}$$

where $u_{i\alpha}(\mathbf{r}-\mathbf{r}_i)$ is the pairwise isotropic interaction potential between solute atom i at position $\mathbf{r}_i$ and solvent site α at $\mathbf{r}$. With the solute-solute forces evaluated directly as $-\partial U_1/\partial \mathbf{r}_i$ and the effect of solvent accounted with mean solvation forces eq. (6), the equations of quasidynamic motion are solved only for solute atoms. In the approach of an adaptive box, the 3D-RISM-KH integral equations eqs. (2) and (3) are discretized and converged on a grid in a non-periodic box of size and shape that includes about 2-3 solvation shells around the biomolecule to minimize boundary effects and varies during the simulation adjusting to solute conformational changes so as to optimize computational load.[35] This is different from conventional MD which typically uses a periodic rectangular box and the Ewald summation technique[109,110] to evaluate long-range electrostatic interactions.

5 Generalized Solvation Force Extrapolation

As mentioned in the Introduction, the original SFE scheme[35] is restricted to small outer time steps, whereas the ASFE method[37] still needs significant modification and generalization to be efficient not just for simple solutes but for large macromolecules as well. Shown below is how 3D-RISM-KH mean solvation forces can be extrapolated in the most general case of solvated protein by using the GSFE method.[38]

5.1 Local non-Eckart like rotational transformations

Let $\mathbf{f}_{i,k}$ be the solvation forces acting on solute sites $i = 1, 2, \ldots, M$ at N previous outer time steps $k = 1, 2, \ldots, N$ for which the 3D-RISM-KH integral equations are converged. The atomic positions at these steps will be denoted by $\mathbf{r}_{i,k}$. The forces $\mathbf{f}_{i,k}$ and positions $\mathbf{r}_{i,k}$ for a given i are ordered in such a way that larger values of k correspond to earlier moments t_k of time, i.e., $t_N < t_{N-1} < \ldots < t_2 < t_1$. The next outer moment is denoted by $t_0 > t_1$. Let $\mathbf{r}_i(t)$ be the current coordinate of atom i at some inner time point t belonging to the interval $]t_1, t_0[$. The total number of these points is equal to $P = (t_0 - t_1)/\Delta t \gg 1$, where Δt is the inner time step. Note also that the actual force $\mathbf{f}_i(t)$ exerted on atom i at time t will depend on the multidimensional vector $\{\mathbf{r}_{1,k}, \mathbf{r}_{2,k}, \ldots, \mathbf{r}_{M,k}\}$ via the relative positions $\mathbf{r}_{ij}(t) = \mathbf{r}_i(t) - \mathbf{r}_j(t)$ of $M - 1$ neighbors j (with $j \neq i$). This follows from the translational invariance of solvation interactions when the total system (solute plus solvent) is arbitrarily shifted as a whole.

One of the main ideas of this approach is to find such local rotational transformations $\mathbf{R}_{ij} = \mathbf{S}_i \mathbf{r}_{ij}$ of the relative positions $\mathbf{r}_{ij}$ for each atom $i = 1, 2, \ldots, M$ (where $j = 1, 2, \ldots, M$) that provide the most smooth behavior of $\mathbf{F}_i = \mathbf{f}_i(\{\mathbf{R}_{ij}\})$ in the new coordinates. This can be achieved by reducing the coordinate region of force extrapolation. For the discrete set ($k = 1, 2, \ldots, N$) of the basis coordinate knots $\mathbf{r}_{i,k}$ the desired transformation $\mathbf{R}_{ij,k} = \mathbf{S}_{i,k}\mathbf{r}_{ij,k}$ with $\mathbf{r}_{ij,k} = \mathbf{r}_{i,k} - \mathbf{r}_{j,k}$ can be determined by minimizing the normalized distances between all the transformed outer coordinates $\mathbf{R}_{ij,k}$ and some origin (where $\mathbf{S} \equiv \mathbf{I}$) point $\mathbf{r}^*_{ij}$ lying in the extrapolating region as

$$\frac{1}{M_i} \sum_{j=1}^{M}{}' w(r^*_{ij}) \Big(\mathbf{S}_{i,k}\mathbf{r}_{ij,k} - \mathbf{r}^*_{ij} \Big)^2 = \min \ . \tag{7}$$

Here w is a weighting function, $\sum'$ stands for $j \neq i$, and M_i is the total number of neighbors for which $w(r^*_{ij}) \neq 0$. The current inner coordinate $\mathbf{r}_{ij}(t)$ should also be transformed analogously by $\mathbf{R}_{ij}(t) = \mathbf{S}_i(t)\mathbf{r}_{ij}(t)$ with

$$\frac{1}{M_i} \sum_{j=1}^{M}{}' w(r^*_{ij}) \Big(\mathbf{S}_i(t)\mathbf{r}_{ij}(t) - \mathbf{r}^*_{ij} \Big)^2 = \min \quad \text{for each } \ i = 1, 2, \ldots, M. \tag{8}$$

Any choice for $\mathbf{r}^*_{ij} = \mathbf{r}_{ij}(t^*)$ with $t_1 \leq t^* \leq t$ can be in principle acceptable, where t is the current inner time and t_1 is the most recent point from the basis outer steps. However, the limiting values $t^* = t_1$ and $t^* = t$ are not recommended in the context of efficiency. Note that in eq. (7) we should carry out the transformation for each $k = 1, 2, \ldots, N$ (and $i = 1, 2, \ldots, M$) whenever $\mathbf{r}^*_{ij}$ is changed, i.e., up NMP times if $t^* = t$, but only NM

ones for $t^* = t_1$. In the latter case, however, the origin $\mathbf{r}^*_{ij}$ equal to $\mathbf{r}_{ij,1}$ may appear to be too far from the current point $\mathbf{r}_{ij}(t)$ when the size of the outer time step $h = t_0 - t_1$ is large. Thus, an optimal choice is when the origin $\mathbf{r}^*_{ij}$ of the transformation is updated after every $1 \ll p \ll P$ inner time step during the outer interval $(t_0 - t_1)$. This constitutes the so-called frequency reuse scheme.

The necessity of introducing the weighting function $w(r_{ij})$ is dictated by the fact that the neighbors with smaller interatomic distances r_{ij} contribute to the mean solvation force $\mathbf{f}_i$ more significantly and thus are more important for the minimization. The most natural way to model such a situation is to put $w(r_{ij}) = 1/r^2_{ij}$, meaning that the relative (and not absolute) interatomic distances are minimized. Then the left-hand sides of eqs. (7) and (8) become dimensionless. At long enough $r_{ij} > R$ the correlations between $\mathbf{r}_{ij}$ and $\mathbf{f}_i$ are diminished, and so we can put $w(r_{ij}) = 0$ in this range. It should be emphasized that the above truncation concerns only the coordinate transformations [eqs. (7) and (8)] but not the actual solvation forces $\mathbf{f}_{i,k}$ which are calculated for all atomic pairs, without any cutoff. Moreover, for small solute molecules of radius less than R, no truncation is performed at all. For large macromolecules with $M \gg 1$, setting a finite cutoff radius R can considerably improve the quality of the mean force extrapolation using only a relatively small number $N \ll M$ of the basis outer points.

The simplest way to obtain explicit expressions for the rotational matrix $\mathbf{S}_{i,k}$ or $\mathbf{S}_i(t)$ is to represent them in terms of the four components quaternion $\mathbf{q} = \{\chi, \eta, \xi, \zeta\}$ as[111]

$$\mathbf{S} = \begin{pmatrix} \chi^2 + \eta^2 - \xi^2 - \zeta^2 & 2(\eta\xi - \chi\zeta) & 2(\chi\xi + \eta\zeta) \\ 2(\chi\zeta + \eta\xi) & \chi^2 + \xi^2 - \eta^2 - \zeta^2 & 2(\xi\zeta - \chi\eta) \\ 2(\eta\zeta - \chi\xi) & 2(\chi\eta + \xi\zeta) & \chi^2 + \zeta^2 - \eta^2 - \xi^2 \end{pmatrix} \tag{9}$$

with $\mathbf{q}^2 \equiv \mathbf{q}^+\mathbf{q} = \chi^2 + \eta^2 + \xi^2 + \zeta^2 = 1$. Inserting now eq. (9) into the superposition equation [eq. (7) or (8)] yields

$$\frac{1}{2}\mathbf{q}^+\mathbf{\Theta}_i\mathbf{q} - \frac{1}{2}\vartheta(\mathbf{q}^+\mathbf{q} - 1) = \min\,, \tag{10}$$

where $\mathbf{\Theta}_i = \frac{1}{M_i}\sum_{j=1}^{M}{}' w(r^*_{ij})\mathbf{\Theta}_{ij}$,

$$\mathbf{\Theta}_{ij} = \begin{pmatrix} (\mathbf{r}'_{ij} - \mathbf{r}^*_{ij})^2 & 2(\mathbf{r}'_{ij} \times \mathbf{r}^*_{ij})^+ \\ 2(\mathbf{r}'_{ij} \times \mathbf{r}^*_{ij}) & \mathbf{I}(\mathbf{r}'_{ij} + \mathbf{r}^*_{ij})^2 - 2(\mathbf{r}'_{ij}\mathbf{r}^{*\,+}_{ij} + \mathbf{r}^*_{ij}\mathbf{r}'^{\,+}_{ij}) \end{pmatrix} \tag{11}$$

are the symmetric 4×4 matrices, $\mathbf{r}'_{ij}$ is equal either to $\mathbf{r}_{ij,k}$ or $\mathbf{r}_{ij}(t)$ for the cases $\mathbf{S}_{i,k}$ or $\mathbf{S}_i(t)$, respectively, ϑ is the Lagrange multiplier, $\mathbf{I}$ is the identity 3×3 matrix, and $\times$ denotes the vector product. Differentiating eq. (10) with respect to all the four components of $\mathbf{q}$ leads to the eigenvalue problem

$$\mathbf{\Theta}_i\mathbf{q} = \vartheta\mathbf{q}\,. \tag{12}$$

Because the right-hand sides of eqs. (7) and (8) are always greater or equal to zero, the matrix $\mathbf{\Theta}_i$ is positive semidefinite, having four eigenvectors $\mathbf{q}_{1,2,3,4}$ and the same number of nonnegative associated eigenvalues $\vartheta_{1,2,3,4} \geq 0$. The latter can be sorted in the ascending order, such that ϑ_1 is the smallest eigenvalue. It coincides with the global minimum

in eqs. (7), (8) and (10) since for any normalized eigenvectors the following equality takes place: $\mathbf{q}^{+}\mathbf{\Theta}_i\mathbf{q} = \vartheta$. The normalized eigenvector $\mathbf{q}_1$ corresponding to the smallest eigenvalue $\vartheta \equiv \vartheta_{1,(i,k)}$ or $\vartheta_{1,i}$ is thus the quaternion describing the optimal transformation by the rotational matrix $\mathbf{S}$ [eq. (9)].

5.2 Individual minimization by weighted least squares

Having the transformed coordinates

$$\mathbf{R}_{ij,k} = \mathbf{S}_{i,k}\mathbf{r}_{ij,k}\,, \qquad \mathbf{R}_{ij}(t) = \mathbf{S}_i(t)\mathbf{r}_{ij}(t)\,, \tag{13}$$

the solvation forces can be extrapolated as follows. First, for each atom i, the actual neighboring positions $\mathbf{R}_{ij}(t)$ are virtually approximated at a given inner point t of the next outer time interval $]t_1, t_0[$ by a linear combination of their previous outer values as

$$\tilde{\mathbf{R}}_{ij}(t) = \sum_{k=1}^{N} A_k^{(i)}(t)\mathbf{R}_{ij,k}\,. \tag{14}$$

The expansion coefficients $A_k^{(i)}(t)$ in eq. (14) can then be obtained as the best representation of the solute neighboring coordinates $\mathbf{R}_{ij}(t)$ at time t in terms of their projections onto the basis of N previous outer positions $\mathbf{R}_{ij,k}$ by minimizing a weighting norm of the difference between $\mathbf{R}_{ij}(t)$ and their approximated counterparts $\tilde{\mathbf{R}}_{ij}(t)$. Additionally imposing the normalizing conditions

$$\sum_{k=1}^{N} A_k^{(i)} = 1\,, \qquad \sum_{k=1}^{N} {A_k^{(i)}}^2 = \min\,, \tag{15}$$

the above minimization leads for each $i = 1, 2, \ldots, M$ to the following modified least-square problem

$$\frac{1}{M_i}\sum_{j=1}^{M}{}' w(R_{ij}^*)\left(\mathbf{R}_{ij} - \sum_{k=1}^{N} A_k^{(i)}\mathbf{R}_{ij,k}\right)^2 + 2\Lambda_i\left(\sum_{k=1}^{N} A_k^{(i)} - 1\right) + \varepsilon_i^2\sum_{k=1}^{N}{A_k^{(i)}}^2 = \min\,, \tag{16}$$

where Λ_i is the Lagrangian multiplier, and $\varepsilon_i^2 \geq 0$ is a balancing parameter. Note that for coordinate deviations the weighting function $w(R_{ij}^*) \equiv w(r_{ij}^*)$ used in eq. (16) is the same as in local rotations [eqs. (7) and (8)], meaning again that neighbors j lying more closely to the reference atom i should be mapped more accurately. Note also that $R_{ij}^* = r_{ij}^*$ because the rotation transformation is unitary ($\mathbf{S} \equiv \mathbf{I}$) in the origin point $\mathbf{r}_{ij}^* \equiv \mathbf{R}_{ij}^*$.

Now, the forces $\mathbf{F}_i(t)$ at any inner time $t \in]t_1, t_0[$ can be extrapolated on the basis of their outer values $\mathbf{F}_{i,k}$ employing a linear expansion procedure which is quite similar to that [eq. (14)] for coordinates $\mathbf{R}_{ij}(t)$. This yields

$$\tilde{\mathbf{F}}_i(t) = \sum_{k=1}^{N} A_k^{(i)}(t)\mathbf{F}_{i,k}\,, \tag{17}$$

where $i = 1, 2, \ldots, M$ and the expansion coefficients $A_k^{(i)}(t)$ are the same as those in eq. (14). This is justified by the fact that $\mathbf{F}_i(t)$ is a function of only $\mathbf{R}_{ij}(t)$. Thus, a better representation of the coordinates by $\tilde{\mathbf{R}}_{ij}(t)$ should provide a more accurate extrapolation

of the interactions, expecting a small difference between the exact forces $\mathbf{F}_i(t)$ and their approximated values $\tilde{\mathbf{F}}_i(t)$. Note that the coordinate mapping is virtual in the sense that $\mathbf{R}_{ij}(t)$ are never replaced by $\tilde{\mathbf{R}}_{ij}(t)$. It is necessary only to find the coefficients $A_k^{(i)}$ for the actual force approximation. Remember also that in eq. (16) the weighting function $w(r)$ being equal to $1/r^2$ for $r \leq R$ is truncated by $w(r) = 0$ at longer $r > R$. Such a truncation in eqs. (7), (8) and (16) may lead to some boundary effects. However, these effects can be neglected by choosing the radius R to be large enough.

An issue now arises how to obtain $\mathbf{F}_{i,k}$ from $\mathbf{f}_{i,k}$ without direct recalculations $\mathbf{F}_{i,k} = \mathbf{f}(\{\mathbf{R}_{ij,k}\})$, and in which way to return back from $\tilde{\mathbf{F}}_i$ to the desired extrapolated forces $\tilde{\mathbf{f}}_i$ in the usual coordinate space. This issue can be solved by taking into account that the original solvation forces $\mathbf{f}_i$ are not only translationally invariant but also satisfy the following orientational condition

$$\mathbf{f}(\{\mathbf{S}_i\mathbf{r}_{ij}\}) = \mathbf{S}_i\mathbf{f}(\{\mathbf{r}_{ij}\})\,, \tag{18}$$

where $\mathbf{S}$ is an arbitrary 3×3 rotational matrix. eq. (18) merely states that if the solute molecule is rotated as a whole, the total solvation forces acting on each atom of this molecule will be transformed according to the same rotation. Note that in this case of the detailed extrapolation, the atomic coordinates are defined relatively to the current reference site $i = 1, 2, \ldots, M$. This means that the virtual rotation of $\mathbf{r}_{ij} \equiv \mathbf{r}_{ij,k}$ or $\mathbf{r}_{ij}(t)$ by $\mathbf{S}_i \equiv \mathbf{S}_{i,k}$ or $\mathbf{S}_i(t)$ is performed in eq. (18) at each given i around this site for all other atoms j. On the other hand, the rotations in eqs. (7) and (8) are performed only for groups of atoms for which $r_{ij}^* \leq R$ with $j \neq i$ and the truncation of neighbors is possible.

From eq. (18) it follows that

$$\mathbf{F}_{i,k} = \mathbf{f}(\{\mathbf{R}_{ij,k}\}) = \mathbf{f}(\{\mathbf{S}_{i,k}\mathbf{r}_{ij,k}\}) = \mathbf{S}_{i,k}\mathbf{f}_{i,k} \tag{19}$$

$$\mathbf{F}_i(t) = \mathbf{f}(\{\mathbf{R}_{ij}(t)\}) = \mathbf{f}(\{\mathbf{S}_i(t)\mathbf{r}_{ij}(t)\}) = \mathbf{S}_i(t)\mathbf{f}_i(t)\,. \tag{20}$$

In view of eqs. (19) and (20), no additional direct recalculations are needed, and the desired approximated forces in the usual coordinate space at each inner point t can be readily reproduced from eq. (17) using the inverse rotational transformation

$$\tilde{\mathbf{f}}_i(t) = \mathbf{S}_i^{-1}(t)\tilde{\mathbf{F}}_i(t) = \mathbf{S}_i^{-1}(t)\sum_{k=1}^{N} A_k^{(i)}(t)\mathbf{S}_{i,k}\mathbf{f}_{i,k}\,. \tag{21}$$

Moreover, the inverse matrix can easily be evaluated taking into account that the rotational transformation is orthonormal, i.e., $\mathbf{S}^{-1} = \mathbf{S}^{+}$, where $\mathbf{S}^{+}$ denotes the transposed matrix.

Therefore, the proposed individual transformation efficiently excludes local rotations of the solute molecule, which can be large enough due to the interactions with the solvent and thermostats. This reduces the volume of the local coordinate space around each solute atom in view of eqs. (7) and (8). Obviously, then a better accuracy of the force extrapolation is provided. In other words, the differences between the approximated values $\tilde{\mathbf{f}}_i(t)$ and their original counterparts $\mathbf{f}_i(t)$ will decrease. In particular, the new extrapolation scheme [eq. (21)] is exact, i.e., $\tilde{\mathbf{f}}_i(t) = \mathbf{f}_i(t)$, already at $N = 1$ for the case of rotating rigid segments constituting the molecule, where the transformed forces $\mathbf{F}_i$ are constant. It should also be very precise for flexible segments, since the magnitudes of the atomic vibrational oscillations are small. The influence of torsion movements on $\mathbf{F}_i$ can also be minimized by extending the basis set and choosing the best subset.

5.3 Normal equations with balancing

The most gentle way to find the coefficients $A_k^{(i)}$ in the force extrapolation [eq. (19)] is to reduce the least-square minimization [eq. (16)] to a normal representation.[112,113] Differentiating eq. (16) with respect to these coefficients and Λ_i leads to the following set of $N+1$ linear equations

$$\begin{pmatrix} G_{11}^{\varepsilon(i)} & G_{12}^{(i)} & \dots & G_{1N}^{(i)} & 1 \\ G_{21}^{(i)} & G_{22}^{\varepsilon(i)} & \dots & G_{2N}^{(i)} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ G_{N1}^{(i)} & G_{N,2}^{(i)} & \dots & G_{NN}^{\varepsilon(i)} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} A_1^{(i)} \\ A_2^{(i)} \\ \vdots \\ A_N^{(i)} \\ \Lambda_i \end{pmatrix} = \begin{pmatrix} G_1^{(i)} \\ G_2^{(i)} \\ \vdots \\ G_N^{(i)} \\ 1 \end{pmatrix} \tag{22}$$

which should be solved for the same number of unknowns $A_k^{(i)}$ at $k = 1, 2, \dots, N$ and Λ_i at each $i = 1, 2, \dots, M$, where $G_{kk}^{\varepsilon(i)} = G_{kk}^{(i)} + \varepsilon_i^2$ with

$$G_{kl}^{(i)} = \frac{1}{M_i} {\sum_{j=1}^{M}}' w(R_{ij}^*) \mathbf{R}_{ij,k} \cdot \mathbf{R}_{ij,l} \,, \tag{23}$$

$$G_{k}^{(i)} = \frac{1}{M_i} {\sum_{j=1}^{M}}' w(R_{ij}^*) \mathbf{R}_{ij,k} \cdot \mathbf{R}_{ij} \,, \tag{24}$$

and $l = 1, 2, \dots, N$. Note that the $(N+1) \times (N+1)$ square matrix in eq. (22) remains to be symmetrical, since the ε_i^2-addition concerns only diagonal elements.

The Lagrange multiplier Λ normalizes the linear equations with the imposed constraint $\sum_k A_k = 1$ [see eq. (15)]. It is necessary to make the extrapolation to be exact for the spatially homogeneous part of the interactions in the transformed space. Indeed, the solvation force $\mathbf{F}_i(t) = \mathbf{F}_i(\mathbf{R}(t))$ can be expanded in the power series of a deviation of the current coordinate vector $\mathbf{R}$ from the origin $\mathbf{R}^* = \mathbf{S}\mathbf{r}^* \equiv \mathbf{r}^*$ at each $i = 1, 2, \dots, M$ as

$$\mathbf{F}_i(\mathbf{R}) = \mathbf{F}_i(\{\mathbf{R}_{ij}^*\}) + \sum_{j=1}^{M} \left. \frac{\partial \mathbf{F}_i}{\partial \mathbf{R}_{ij}} \right|_{\mathbf{R}_{ij}^*} \left(\mathbf{R}_{ij} - \mathbf{R}_{ij}^*\right), \tag{25}$$

where $\partial \mathbf{F}_i / \partial \mathbf{R}_{ij}$ is the Hessian $(3M \times 3M)$ matrix, and the second and higher order spatial inhomogeneities $\mathcal{O}[(\mathbf{R}_{ij} - \mathbf{R}_{ij}^*)^2]$ have been neglected. Thus $\mathbf{F}_i(\mathbf{R})$ has the constant zeroth-order part representing by the first term in the right-hand side of eq. (25). It immediately follows from eq. (17) that this term can be reproduced exactly, provided $\sum_k A_k = 1$. The second term in the right-hand side eq. (25) is linear in coordinates. That is why it can be extrapolated using the duplex linear expansions [eqs. (14) and (17)]. In fact, each of the $3M \times 3M \gg 1$ elements of the Hessian matrix is mapped in a very complicated way via the obtained solutions for the extrapolation coefficients A_k involving a finite number $(N \gg 1)$ of the basis knots $\mathbf{R}_{ij,k}$.

The balancing parameter $\varepsilon^2 > 0$ appears as a result of the required minimization $\sum_k A_k^2 = \min$ for the norm of the expansion coefficients [see eq. (15)]. Such an additional minimization is also needed for the following reason. The used dual (virtual coordinate

and actual force) extrapolation tentatively assumes that lowering of the coordinate residuals should immediately lead to a decrease of the deviations between the approximated and original forces. But this is not so when N approaches the number $3M_i$ of local internal degrees of freedom of the neighbor atoms. Then the least-square solver to eq. (16) will try to reduce the coordinate residuals to the global (zeroth) minimum with no regarding to the values of expansion coefficients A_k which are exploited in both the coordinate and force extrapolations. As a result, a lot of these coefficients may accept large negative and positive values, despite the presence of the linear normalizing condition $\sum_k A_k = 1$. It is well known from the general theory of extrapolative and quadrature formulas that the existence of weights large in magnitude decreases the stability range, leading to an appreciable increase of the uncertainties outside of this region.

The minimization $\sum_k A_k^2 = \min$ is introduced just to avoid the above singularity at $N \sim 3M_i$ when $\varepsilon = 0$. The non-zero values of $\varepsilon^2 > 0$ allow us to effectively balance between the two kinds of the extrapolations. Of course, ε^2 cannot be chosen too large because then the main effort is directed to minimizing the squared norm $\varepsilon^2 \sum_k A_k^2 = \min$ rather than the coordinate residuals. This parameter should be treated as a small quantity aiming at improving the quality of solvation force extrapolation. Optimal values of ε^2 can be found in actual simulations to obtain the best accuracy.

5.4 Extending the basis set and selecting the best subset

Evidently, the accuracy of the force extrapolation should increase with increasing the number N of basis points. However, N cannot be put too large because then the number of linear equations increases, too. These equations [eq. (22)] need to be solved frequently (in total $P = h/\Delta t \gg 1$ times per h), namely, at each inner point inside the outer interval $h \gg \Delta t$ for the minimization of coordinate residuals [eq. (16)]. As a result, the computational overhead can be unacceptably high at large enough values of N, reducing the efficiency of MD/3D-RISM-KH simulation.

A way to remedy the above situation lies in the following. The basis set can be extended from a relatively small number of $N \lesssim 100$, say, to a larger value $N' \gg N$ by collecting the force-coordinate pairs during a broad previous time interval $\Delta H = N'h \gg Nh$. Then the weighting squared distances in the $3M$-dimensional space between the transformed basis outer coordinates $\mathbf{R}_{ij,k'}$ and the current origin point $\mathbf{r}_{ij}^*$,

$$\mathcal{R}_{ik'}^2(t) = \frac{1}{M_i} {\sum_{j=1}^{M}}' w(r_{ij}^*) \left(\mathbf{R}_{ij,k'} - \mathbf{r}_{ij}^*\right)^2 \equiv \vartheta_{1,(i,k')} \,, \tag{26}$$

can be readily expressed for each $i = 1, 2, \ldots, M$ and $k' = 1, 2, \ldots, N'$ in terms of the smallest eigenvalues $\vartheta_{1,(i,k')}$ [see the text after eq. (12)]. Now these distances can be sorted in the ascending order, and the first N most closest points can be selected among the extended set to satisfy the condition $\mathcal{R}_{i1} < \mathcal{R}_{i2} < \ldots < \mathcal{R}_{iN}$. The forces $\mathbf{F}_{i,k'}$ must be resorted synchronically with the coordinates $\mathbf{R}_{ij,k'}$ to form the best pair subset with N points. It should then be used when performing the advanced extrapolation [eq. (21)].

The above procedure can further improve the quality of the extrapolation, especially at $N' \gg N$. The reason is that the choice of the nearest outer pairs in the transformed space additionally reduces the coordinate region in which the extrapolation is performed. This

leads to a decrease of the coordinate residuals and, as a consequence, to an increase of the accuracy. In fact, such an additional reduction minimizes the change in the transformed solvation forces during torsion motion of the solute. Note that such motion (characterizing by large amplitudes) is responsible for transitions of the biomolecule from one conformational pool to another where the torsion potential has a local minima. Thus, an optimal value for the expanded interval $\Delta H = N'h$ should be of order of the mean life time in local conformational minima. Then, whenever the transition to other conformations occurs, the subset can be quickly reselected to fit the basis outer points to the current solute conformation. The accuracy of such fitting is especially high if the molecule has already been near this conformation at previous times.

It is worth remarking that the selecting procedure at $N \ll N'$ requires only little extra numerical efforts even for large enough N' (of order of several thousands). This is explained by the fact that computational cost grows with increasing N' just slightly. Indeed, the selection operates only with the smallest eigenvalues $\vartheta_{1,(i,k')}$ (and not with eigenvectors) of small 4×4 matrices, and so the computational time is linearly proportional to N' (with a small coefficient near N'). The eigenvectors $\mathbf{q}_{1,(i,k)}$ are necessary only for the best subset with $k = 1, 2, \ldots, N \ll N'$ to build the transformation matrix $\mathbf{S}_{i,k}$ for the extrapolation [eq. (21)]. On the other hand, the overhead increases much more rapidly with increasing N, namely, proportionally up to $(N+1)^3$, as this is required to find solutions to $(N+1)$ linear equations [eq. (22)]. In addition, the selection procedure is performed only once per many ($p \gg 1$) inner time steps, further lowering the computational costs.

5.5 The whole algorithm of GSFE

With the techniques laid out in the preceding subsections, the resulting generalized solvation force extrapolation (GSFE) algorithm can be briefly described as follows.

At the very beginning, the 3D-RISM-KH integral equations are converged after each Δt of the N first inner steps with no extrapolation to fill out the basis set. Then the extrapolation starts with N points and the extended N'-set is accordingly completed step by step in the integration process. Since h can be much larger than Δt, the outer step cannot be put immediately equal to $h \gg \Delta t$. The reason is that then the extrapolation skews because of the significant non-uniformity of the time intervals between the points from the set. This issue can be remedied in such as a way that the outer time interval is smoothly increased every inner step from Δt to h with an increment of Δt.

Further, after each $p\Delta t$ step, the eigenvalue problem [eqs. (11) and (12)] is solved for the extended set with N' coordinates. The first $N < N'$ points are selected by sorting the corresponding smallest eigenvalues [eq. (26)] in the ascending order. The coordinates and forces related to the subset obtained are then transformed by the local non-Eckart like rotations [eq. (19)] in terms of the $\mathbf{S}$-matrix [eq. (9)] constructed on the N eigenvectors. Having the transformed coordinates, the system of $(N+1)$ linear equations [eq. (22)] is built and solved for the expansion coefficients. Note that the inversion of the $(N+1)\times(N+1)$ matrix in [eq. (22)] is performed only once per p inner steps because it remains unchanged during time $p\Delta t$ [see eq. (23)], while the right-hand side vector in [eq. (22)] varies [eq. (24)] every Δt.

Using the expansion coefficients, the solvent forces are extrapolated at each inner step Δt within the outer time interval $t \in]t_1, t_0[$ of length h as the weighted sum of their N

previous outer transformed values, followed by the inverse transformation [eq. (21)]. The extrapolation procedure is applied $h/\Delta t$ times to achieve the next outer point. At that point, the solvent forces are calculated explicitly by solving the 3D-RISM-KH integral equations. The extended N'-set is then updated by the new outer force-coordinate pair, while the oldest one is discarded. All these actions are repeated H/h times for the next outer intervals until the desired simulation time length H is achieved.

This completes the derivation of the GSFE algorithm applicable to arbitrary biomolecules.[38] Formal replacement $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ with $\mathbf{r}_i - \mathbf{r}_c$ and $\mathbf{r}^*_{ij}$ with $\mathbf{r}_{i,1}$, where $\mathbf{r}_c$ is the center of mass of the solute molecule as well as putting $w \equiv 1$ with $R \to \infty$ and $M_i \equiv M$ leads to the global non-Eckart transformation used in ASFE.[37] For large molecules, however, the global rotations appear to be small, thus having practically no effect on extrapolation improvement. Moreover, the extended set and the inversion of the $(N+1)\times(N+1)$ matrix in ASFE are carried out each ($p = 1$) inner time step, which significantly lowers the efficiency of the computations. In GSFE, applying the frequency regime with $1 \ll p \ll P = h/\Delta t$ appreciably reduces computational cost. This is achieved without loss of precision since the increased distances during time $p\Delta t$ are compensated by the inverse non-Eckart-like local transformations.

Worth pointing out is also that the rotational superpositions [eqs. (7) and (8)] introduced for GSFE look somewhat similar to those originally derived for analysis of macromolecular structures obtained in conventional MD or experiment.[114] Note that there is no unique approach to separate translational, angular, and internal motions of these molecules. Within the well-known Eckart scheme[115–120] such a separation can be carried out unambiguously only for molecular structures with one equilibrium state. It does not work for more complicated molecules where two or more local equilibrium states can exist. In Ref.[114], this problem was solved exploiting Gauss' principle of least constrain by minimizing the coordinate deviation norm $\frac{1}{M}\sum_{i=1}^{M} m_i\big(\delta\mathbf{S}(t,\delta t)\mathbf{r}_i(t+\delta t) - \mathbf{r}_i(t)\big)^2 = \min$, where m_i is the mass of ith atom and δt is the time step. For instance, in the limit $\delta t \to 0$, this allows us to uniquely define the angular velocity $\mathbf{\Omega}(t) = \mathbf{1}(t)\lim_{\delta t\to 0}\delta\phi/\delta t$ of an arbitrary molecule at any time t, where $\delta\mathbf{S}(t,\delta t) \equiv \mathbf{S}(\mathbf{\Delta q}(t))$ with $\mathbf{\Delta q}(t) = \{\cos(\delta\phi/2), \mathbf{1}(t)\sin(\delta\phi/2)\}$ is the matrix of rotation of the whole molecule by angle ϕ around the unit vector $\mathbf{1}$ passing through the center of mass. The standard Eckart method appears as a particular case of the non-Eckart approach when the number of local equilibrium states is equal to one. Moreover, in the absence of internal degrees of freedom, the non-Eckart angular velocity completely coincides with the well known definition for rigid bodies.

The present non-Eckart like superposition scheme differs in several aspects from the original non-Eckart method.[114] It is modified by normalizing weights and applied individually for each reference atom of the solute molecule at discrete moments of time. This results in local reorientations of atomic groups instead of those in rotation of the molecule as a whole. Moreover, rather than only analyzing simulation or experimental data,[114,121–123] the newly introduced non-Eckart like scheme is aimed at optimizing the performance of MD simulations.

6 OIN Ensemble MTS-MD Steered with Extrapolated 3D-RISM-KH Mean Solvation Forces

The equations of motion for solute atoms in hybrid MD/3D-RISM-KH simulation in the canonical-isokinetic OIN ensemble steered with 3D-RISM-KH mean solvation forces which are extrapolated with the GSFE technique can be cast in the compact form[36]

$$\frac{d\mathbf{\Gamma}}{dt} = L\mathbf{\Gamma}(t)\,, \tag{27}$$

where $\mathbf{\Gamma} = \{\mathbf{r}, \mathbf{v}; \boldsymbol{\varsigma}, \mathbf{w}\}$ denotes the extended phase space and L is the Liouville operator. The extended space, apart from the full set of coordinates $\mathbf{r} \equiv \{\mathbf{r}_i\}$ and velocities $\mathbf{v} \equiv \{\mathbf{v}_i\}$ of all solute atoms, includes also all thermostat frequencies $\boldsymbol{\omega} \equiv \{\omega_{\kappa,i}\}$ with $\kappa = 1, \ldots, \mathcal{K}$ and their conjugated dynamical variables $\boldsymbol{\varsigma} \equiv \{\varsigma_i\}$. The latter are introduced by means of the relation $d\varsigma_i/dt = (\tau_i^2\omega_{1,i}^2\omega_{2,i} - \sum_{\kappa=2}^{\mathcal{K}}\omega_{\kappa,i})$, where $\mathcal{K}$ is the number of chains per thermostat. The Liouvillian can be split up as

$$L = \sum_{i=1}^{M}\left(\mathcal{A}_i + B_i + C_{v,\omega,i} + C_{\omega,i} + C_{\varsigma,i}\right) \tag{28}$$

into the kinetic $\mathcal{A}_i = \mathbf{v}_i \cdot \partial/\partial\mathbf{r}_i$, potential

$$B_i = \left(\frac{\boldsymbol{f}_i}{m_i} - \mathbf{v}_i\frac{\mathbf{v}_i \cdot \boldsymbol{f}_i}{2\mathrm{T}_i}\right) \cdot \frac{\partial}{\partial\mathbf{v}_i} - \frac{\mathbf{v}_i \cdot \boldsymbol{f}_i}{2\mathrm{T}_i}\omega_{1,i}\frac{\partial}{\partial\omega_{1,i}}\,, \tag{29}$$

and chain-thermostat parts

$$C_{v,\omega,i} = \frac{\tau_i^2\omega_{1,i}^2}{4}\omega_{2,i}\mathbf{v}_i \cdot \frac{\partial}{\partial\mathbf{v}_i} + \left(\frac{\tau_i^2\omega_{1,i}^2}{4} - 1\right)\omega_{1,i}\omega_{2,i}\frac{\partial}{\partial\omega_{1,i}}\,, \tag{30}$$

$$C_{\omega,i} = \sum_{\kappa=2}^{\mathcal{K}}\left(\omega_{\kappa-1,i}^2 - \frac{1}{\tau_i^2} - \omega_{\kappa+1,i}\omega_{\kappa,i}\right)\frac{\partial}{\partial\omega_{\kappa,i}}\,, \tag{31}$$

$$C_{\varsigma,i} = -\left(\tau_i^2\omega_{1,i}^2\omega_{2,i} - \sum_{\kappa=2}^{\mathcal{K}}\omega_{\kappa,i}\right)\frac{\partial}{\partial\varsigma_i}\,. \tag{32}$$

In the canonical OIN ensemble[36] each atom is coupled with its own thermostat by imposing the constraint $\mathrm{T}_i = 3k_\mathrm{B}T/2$, where

$$\mathrm{T}_i = \frac{m_i\mathbf{v}_i^2}{2} + \frac{3k_\mathrm{B}T}{4}\frac{\tau_i^2\omega_{1,i}^2}{2} \tag{33}$$

is the full kinetic energy of the i-th subsystem. The quantity τ_i is related to the relaxation time, determining the strength of coupling of atom i with its thermostat.

The total forces $\boldsymbol{f}_i = \mathbf{f}_{i(\mathrm{f})} + \mathbf{f}_i$ are now divided into the fast solute-solute component $\mathbf{f}_{i(\mathrm{f})}$ and slow 3D-RISM-KH solute-solvent one $\mathbf{f}_i$. In view of eq. (29), this results in the corresponding splitting of the potential operator as $B_i(\{\boldsymbol{f}_i\}) = B_i(\{\mathbf{f}_{i(\mathrm{f})}\}) + B_i(\{\mathbf{f}_i\}) \equiv B_\mathrm{f} + B_\mathrm{s}$. Mention that the solute-solute forces $\mathbf{f}_{i(\mathrm{f})}$ are calculated always directly (by $-\partial U_1/\partial\mathbf{r}_i$), while the 3D-RISM-KH solute-solvent mean forces $\mathbf{f}_i$ are either evaluated

explicitly in the form eq. (6) or approximated with $\tilde{\mathbf{f}}_i$ using the transformation eq. (21). Then $B_{\rm s}(\{\mathbf{f}_i\})$ transforms to $B_{\rm s}(\{\tilde{\mathbf{f}}_i\}) \equiv \tilde{B}_{\rm s}$.

Thus, using the MTS decomposition method,[60,66,67,120] the solution $\mathbf{\Gamma}(h) = e^{Lh}\mathbf{\Gamma}(0)$ to eq. (27) over the outer time interval h from an initial state $\mathbf{\Gamma}(0)$ can be presented[36] as the following product of exponential operators:

$$\mathbf{\Gamma}(h) = \prod_{n'=1}^{n} e^{C\frac{\delta t}{2}} e^{B_{\rm fs}^{(n')}\frac{\delta t}{2}} e^{\mathcal{A}\delta t} e^{B_{\rm fs}^{(n')}\frac{\delta t}{2}} e^{C\frac{\delta t}{2}} \mathbf{\Gamma}(0) + \mathcal{O}(\delta t^2)\,. \tag{34}$$

Here, $n = h/\delta t \gg 1$ is the total number of sub-inner time steps with length $\delta t \ll \Delta t$ each, $C = C_{v,\omega} + C_{\omega} + C_{\varsigma}$,

$$e^{B_{\rm fs}^{(n')}\frac{\delta t}{2}} = \begin{cases} e^{B_{\rm f}\frac{\delta t}{2}} e^{B_{\rm s}\frac{\Delta t}{2}} & \text{only once per } h \text{ when } n' = 1 \\ e^{B_{\rm f}\frac{\delta t}{2}} e^{\tilde{B}_{\rm s}\frac{\Delta t}{2}} & \text{every } \Delta t/\delta t \text{ inner step } n' \\ e^{B_{\rm f}\frac{\delta t}{2}} & \text{for all other } n' \end{cases} \tag{35}$$

is the generalized velocity propagator, Δt is the inner ($\delta t \ll \Delta t \ll h$) time step, $\mathcal{O}(\delta t^2)$ is the accuracy of the decomposition, and the subscript i is omitted for the sake of simplicity. Note that we should first update (by $e^{C\delta t/2}$ and $e^{B\delta t/2}$) the complete set of velocities $\mathbf{v}_i$ and frequencies $\omega_{\kappa,i}$ belonging to all atoms ($i = 1, 2, \ldots, M$) and thermostat chains ($\kappa = 1, \ldots, \mathcal{K}$) before to change the coordinates $\mathbf{r}_i$ of all particles by $e^{\mathcal{A}\delta t}$. A nice feature of the OIN decomposition is that the action of all the single-exponential operators which arise in eqs. (34) and (35) on $\mathbf{\Gamma}$ can be handled analytically using elementary functions.[36]

Therefore, the propagation $\mathbf{\Gamma}(t) = [\mathbf{\Gamma}(h)]^{t/h}$ of dynamical variables from their initial values $\mathbf{\Gamma}(0)$ to arbitrary future time t can be performed by consecutively applying the single exponential transformations of a phase space point $\mathbf{\Gamma}$ in the order defined in eqs. (34) and (35). As can be seen, the fastest $B_{\rm f}$-component of motion is integrated most frequently, namely, $n = h/\delta t$ times per outer interval h with the smallest (sub-inner) time step δt, while the (original or approximated) slow 3D-RISM-KH forces are applied impulsively only every $\Delta t/\delta t$ sub-inner step, i.e., $h/\Delta t < n$ times. Note that almost all these impulses (when $\Delta t \ll h$) are obtained by employing the extrapolated 3D-RISM-KH forces [eq. (21)] in terms of the operator $\tilde{B}_{\rm s}$, and the explicit 3D-RISM-KH calculations [eq. (6)] are used in $B_{\rm s}$ only once per outer time interval h. Taking into account that the solute-solute forces are much cheaper to evaluate than the solute-solvent ones, obvious speedup is achieved compared to single time step propagation ($n = 1$, $\delta t = h$) without extrapolation ($\Delta t = h$). Moreover, the existence of the impulsive inner time steps of length $\Delta t > \delta t$ gives a possibility of reducing the number of (either extrapolative or direct) 3D-RISM-KH evaluations from $h/\delta t$ to $h/\Delta t$. Finally, applying the GSFE approach allows further significant improvement of the overall efficiency, since 3D-RISM-KH calculations which are the most expensive are performed just once per outer step h.

It is worth emphasizing that the presence of OIN thermostatting terms in B and C [see eqs. (29) to (32)] eliminates MTS resonance instabilities. The latter appear in conventional MD simulation (in the microcanonical and canonical ensembles) already at relatively small values of outer time steps. In hybrid MD/3D-RISM-KH simulation, additional instability can come from uncertainties caused by the approximate character of solvent force extrapolation. The canonical-isokinetic OIN propagation [eq. (34)] efficiently eliminates both these types of instabilities by imposing the individual isokinetic constraint [eq. (33)] on

each solute atom. Moreover, the accurate GSFE approach makes possible to drastically increase the sizes of inner and especially outer time steps compared to standard integration.

In view of the above, the following hierarchy of time steps,

$$\delta t \ll \Delta t \ll h \ll Nh \ll N'h = \Delta H \ll H, \tag{36}$$

should be set in order to achieve optimal performance of hybrid MTS-MD/OIN/3D-RISM-KH simulation using the generalized solvation force extrapolation (GSFE) approach. We thus have up to five time scales. This accomplishes coupling of the generalized solvation force extrapolation (GSFE) with the MTS-MD/OIN/3D-RISM-KH method. The resulting scheme will be referred to as hybrid MTS-MD/OIN/GSFE/3D-RISM-KH simulation, or OIN/GSFE/3D-RISM-KH integrator.

Strictly speaking, quasidynamic behavior obtained in MTS-MD/OIN/GSFE/3D-RISM-KH simulation differs from true dynamics in conventional MD with explicit solvent. In particular, such quasidynamics does not obey the Maxwell velocity distribution, and thus, unlike microcanonical MD, cannot get us real time correlation functions. However, as was proven rigorously,[36] the configurational part of the extended partition function obtained in hybrid MTS-MD/OIN/GSFE/3D-RISM-KH simulation at targeted temperature T does coincide with the true canonical distribution of the physical system in coordinate space. This is a very important feature because the original conformational properties, including spatial atom-atom density distribution functions of the solute macromolecule (or supramolecule), are readily reproduced in MTS-MD/OIN/GSFE/3D-RISM-KH simulation. Such quasidynamic sampling appears to be much more efficient than that following from "real-time" (microcanonical or canonical) MD simulations.

7 Illustrations of the Hybrid MTS-MD/OIN/GSFE/3D-RISM-KH Simulation Method on Biomolecular Problems

The new method of hybrid MTS-MD with OIN thermostat steered with 3D-RISM-KH mean solvation forces using GSFE has been implemented in the Amber 11 package[124] (Assisted Model Building with Energy Refinement). The source files `SANDER.F`, `RUNMD.F`, `MDREAD.F`, `MD.H`, `AMBER_RISM_INTERFACE.F`, and `FCE_C.F` from the original package were modified to run hybrid OIN/GSFE/3D-RISM-KH simulation in parallel by the compiled module `SANDER.RISM.MPI`. The first two Fortran modules were modified to implement a program code for solving the canonical-isokinetic OIN equations of motion by adapting the velocity-Verlet like version of the decomposition integration eq. (34) to its leapfrog-like counterpart. (The differences between the standard velocity-Verlet and leapfrog schemes can be found, for instance, in Refs.[111,125]) The third and fourth modules were altered to organize the input/output for new parameters and observable quantities. The last two Fortran documents were rewritten to implement the GSFE procedures.

The accuracy of the OIN/GSFE/3D-RISM-KH integrator, including generalized solvation force extrapolation and optimized isokinetic Nosé-Hoover chain thermostatting, in reproducing the equilibrium and conformational properties has been validated against explicit solvent MD for solutes of different complexity and stiffness in ambient solvents: fully flexible models of alanine dipeptide (22 atoms) in water, asphaltene dimer (336 atoms) in toluene, miniprotein 1L2Y (304 atoms) and protein G (862 atoms) in water.[38] Details of the hybrid simulation setup and the force fields used are listed in Ref.[38]

7.1 Conformational properties of biomolecules

The hybrid OIN/GSFE/3D-RISM-KH simulations of miniprotein 1L2Y and protein G in ambient water at temperature $T = 300$ K and density 1 g/cm^3 started from the folded crystal conformations obtained in NMR experiment, taken from PDB (protein data bank) structures 1L2Y[126] and 1P7E,[127] respectively. The total duration of quasidynamics was 25 ns for miniprotein 1L2Y and 60 ns for protein G. The conformational properties were checked in terms of atomic root-mean-square deviations (RMSD), radius of gyration $R_{\rm g}$, and tertiary structure of the biomolecules. The RMSD was calculated with respect to the initial conformations, with account for only the most massive atoms, C_α, whereas the radius of gyration was computed for all atoms of the proteins.

Figures 1 and 2 present the tertiary structures of hydrated miniprotein 1L2Y and protein G obtained on OIN/GSFE/3D-RISM-KH quasidynamics of duration $t = 25$ ns with outer time step $h = 1$ ps (left panels) in comparison to their initial crystal structure conformations.[126,127] As is seen, the OIN/GSFE/3D-RISM-KH quasidynamics maintains the secondary and tertiary structures of the proteins very well even at the huge 1 ps outer time step. It is an excellent result, taking into account how many numerical techniques and approximations are involved in this hybrid approach. Note that the main sources of possible uncertainties here are the approximate characters of the force fields, mean solvation force extrapolation, and the 3D-RISM integral equation with the KH closure approximation. A comparison of the tertiary structures shows that these uncertainties do not affect the method ability to reproduce the true conformational behavior. Some difference between the initial and final conformations in Figures 1 and 2 is explained by the fact that they correspond to crystal and liquid states. Moreover, the conventional MD simulations with explicit solvent[38] give the folded conformations very similar to those shown in the right-hand parts of Figures 1 and 2.

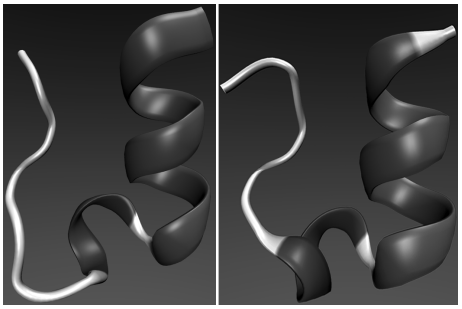


Figure 1: Tertiary structure of miniprotein 1L2Y from experiment[126] before the simulation (left part) and after OIN/GEFE/3D-RISM-KH quasidynamics of duration 25 ns with outer time step 1 ps (right part).

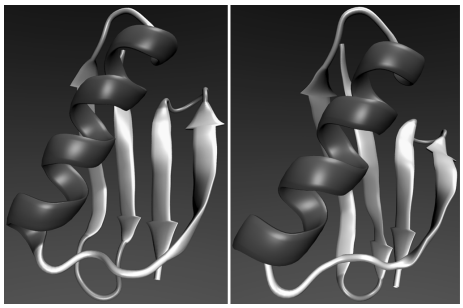


Figure 2: Same as in Figure 1 but for protein G.[127]

7.2 Simulation speedup

The SANDER module in which the MTS-MD/OIN/GSFE/3D-RISM-KH algorithm has been implemented is relatively fast for single CPU core run but not the the best for parallel execution, compared to the PMEMD module, top parallel performer in the Amber MD package. However, SANDER has been available in the GPL version from Amber Tools, and thus presents a convenient platform for testing the method, including scaling up in parallelization. All the simulations were carried out with up to 48 CPU cores per parallel job with 4x Infiniband Interconnect. For the hybrid OIN/GSFE/3D-RISM-KH simulation with 1 ps (and other values of) outer time step for hydrated protein G as well as hydrated miniprotein 1L2Y and asphaltene dimer in toluene, the relative speedup with the number of CPU cores utilized in parallel constitutes 80% to 60% for 4 to 8 CPU cores and then levels out to 50% for up to 48 CPU cores – a typical picture for the interprocessor communication toll in the SANDER module. The efficiency staying at 50% is far from the saturation regime, and so utilizing much more CPU cores in parallel is possible.

The productivity achieved for hydrated protein G in 48 CPU cores parallel run of the OIN/GSFE/3D-RISM-KH quasidynamics with outer time step at the upper limit $h = 1$ ps was 10 ns/day. (WIth SFE instead of GSFE, the upper limit of outer time step drops to $h = 24$ fs and the productivity goes down to 0.44 ns/day; without solvation force extrapolation, to 24 fs and 0.15 ns/day.) The productivity saturates for outer time step increasing to $h \sim 2 - 4$ ps, as the portion of computing effort spent on converging the 3D-RISM-KH integral equations drops inversely with h (about 88;75;50% for $h = 1; 2; 4$ ps) and for $h = 4$ ps goes below the expenses invariably required at each inner time step for calculation of intramolecular solute atomic forces, propagation of coordinates and velocities, and solvation force extrapolation.

The productivity of conventional MD simulation of protein G in explicit water was about 1 ns/day on 48 CPU cores in parallel using the standard SANDER module and all the same simulation setup parameters like the cutoff radii for the solute-solvent interactions and so on. The OIN/GSFE/3D-RISM-KH quasidynamics of hydrated protein G thus achieved actual 10-fold acceleration in terms of a "direct" comparison of the number/time of inner time steps of protein conformation evolution to that in explicit solvent MD. As demonstrated in Section 7.3 below, quasidynamics steered with 3D-RISM-KH mean solvation forces provides 5- to 100-fold time scale compression of protein conformational changes coupled with solvent exchange, thus achieving a huge effective speedup of protein conformational sampling by a factor of 50 to 1000 times compared to conventional MD with explicit solvent or real time dynamics.

7.3 Protein folding

Illustrated below is the of the ability of hybrid MTS-MD/OIN/GSFE/3D-RISM-KH integrator algorithm to produce self-organized, native conformations of proteins, starting from denatured states. Miniprotein 1L2Y was designed more than ten years ago[126] and up to date is the smallest protein to display folding properties. Its 304 atoms constitute a 20-residue amino acid sequence within the so-called tryptophan (Trp) cage (TC5b, PDB code: 1L2Y). The small size and stability of miniprotein 1L2Y at room temperature make it an ideal candidate for computer simulation tests.

So far, there have been no results on folding of this simplest 1l2Y protein by conventional (unbiased) MD simulations in explicit solvent. The reason is that a time interval of 4 to 10 μs is still practically unreachable in one MD run even for modern supercomputers. There are a lot of challenges hindering protein folding simulations.[5] The main difficulty is that protein energy landscape is characterized by a vast number of local minima separated by energy barriers hard to overcome. One of the ways to obviate this problem lies in applying biased replica exchange MD simulation.[128–130] In the replica exchange approach, a great number of short simulations (replicas) are performed in parallel at different temperatures. After certain periods, the conformations are exchanged with a Metropolis probability. As a result, the necessary simulation length corresponding to each replica may be much shorter than the real folding time. However, the whole simulation must span over a wide temperature range with levels spaced closely enough to enable exchanges with high acceptance ratios. This significantly increases the total computational load.

Of much less challenge, high-temperature unfolding MD simulations of the Trp cage in explicit water has been performed.[131] Replica exchange MD simulations of reversible unfolding/folding of the miniprotein in explicit solvent have been done, too.[129,130]

An efficient way to overcome the difficulties inherent in conventional MD of biomolecules is to contract detailed degrees of freedom of individual solvent molecules and perform quasidynamics of the biomolecule steered with mean solvation forces,[68,69] as described in Sec.2. Many works have been devoted to study folding and unfolding pathways in the Trp-cage system using MD simulations with GBSA implicit solvation. Simmerling et al.[132] obtained the folded conformation of miniprotein 1L2Y beginning from an unfolded state. Laser temperature jump relaxation experiments have shown that it is the most rapidly folding protein known with a folding time of order of 4 μs.[133] These experiments were supplemented with stochastic MD/GBSA simulations of the kinetics that indicated a folding time between 1.5 and 8.7 μs.[134] Replica exchange MD in implicit solvent has also been carried out to determine the folding thermodynamics of the miniprotein.[128] Quite recently, the experimentally suggested intermediate and unfolded states in the folding pathway of the Trp-cage miniprotein have been identified in enhanced MD/GBSA simulations.[135]

With the high efficiency of OIN/GSFE/3D-RISM-KH quasidynamics, it has been possible to completely fold miniprotein 1L2Y in 60 ns simulation with 1 ps outer time step, starting from a well denatured, almost fully extended conformation. Similarly to the MD/GBSA simulation,[132] the solution temperature was increased from $T = 300$ to 325 K so as to make the folding time less dependent on initial conditions.

Figure 3 presents four conformational states of hydrated miniprotein 1L2Y obtained on 15, 18, 20, and 60 ns of the OIN/GSFE/3D-RISM-KH quasidynamics starting from the denatured state. As is seen, the miniprotein remains almost completely unfolded like in the initial denatured conformation until 15 ns of simulation, and after that, drastic conformational changes appear quickly. An α-helix begins forming from 18 ns on and expands to its nearly full size by 20 ns of quasidynamics. Furthermore, the 3_{10}-helix secondary structure arises. This looks very similar to the miniprotein native folded states (Figure 1). Some difference apparently means that the simulation time length of 20 ns is still not long enough to achieve the lowest energy state. During the next 40 ns of quasidynamics totaling to the 60 ns long simulation, the miniprotein partially unfolded and folded up again several times but each time in a somewhat different way, confirming the conclusions

from MD/GBSA simulations.[135] Finally, by 60 ns of the quasidynamics, the most folded lowest energy conformation of hydrated miniprotein 1L2Y has been attained. This demonstrates that a protein folding event can be achieved with the unbiased approach of hybrid OIN/GSFE/3D-RISM-KH quasidynamics. (That was previously considered feasible rather with biased methods like replica exchange MD).

Mention that no folding activity was observed in the 60 ns conventional MD simulation of miniprotein 1L2Y in explicit water (with the analogous simulation setup). This is quite reasonable, as in real experiment the miniprotein folds in 4 to 9 μs,[133,134] and so the same folding time scale should be expected in explicit water MD simulation. With the miniprotein folding observed in 60 ns quasidynamics of hybrid OIN/GSFE/3D-RISM-KH simulation, we thus can claim time scale compression by a factor of 100 compared to conventional MD with explicit solvent. (Note that the with the earlier versions of solvation force extrapolation, the time scale compression obtained for hydrated alanine dipeptide was only about 5 times.[37])

With the 10-fold acceleration of time rate of protein conformational evolution compared to that in explicit solvent MD as benchmarked in Section 7.2, the effective speedup of conformational sampling with OIN/GSFE/3D-RISM-KH quasidynamics overall comes to 50- to 1000-fold compared to conventional MD simulation with explicit solvent. The intrinsic acceleration inherent in OIN/GSFE/3D-RISM-KH quasidynamics is expected to further increase with the complexity of proteins since 3D-RISM-KH mean solvation forces properly account for effects of chemical specificities of both solute and solvent molecules in protein confinement, as discussed in Sections 3 and 2.

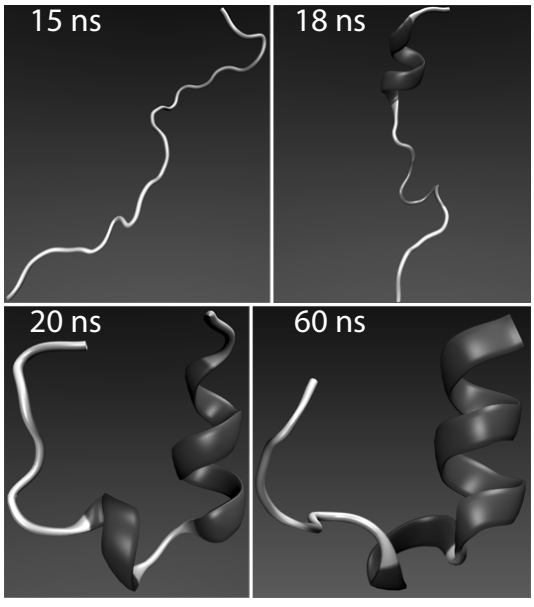


Figure 3: Conformational states of hydrated miniprotein 1L2Y in the OIN/GSFE/3D-RISM-KH folding simulation of duration 25 ns with outer time step 1 ps.

The hybrid MTS-MD/OIN/GSFE/3D-RISM-KH integrator can efficiently sample phase space of biomolecules for essential solvation events with rare statistics such as protein conformational changes coupled with solvation exchange and localization. This provides a substantial gain over conventional MD with explicit solvent requiring an enormous number of time steps and computational time in such cases. In this context, the use of the 3D-RISM-KH molecular theory of solvation appears to be similar to some extent to other techniques enhancing statistical convergence, such as replica exchange MD, umbrella sampling, weighted histogram, parallel tempering, and other biased methods[22] that were originally developed for conventional MD or Monte Carlo simulations. An advantage of MTS-MD/OIN/GSFE/3D-RISM-KH quasidynamics is that it can dramatically improve conformational sampling with much lower computational cost.

8 Concluding Remarks

In this hybrid simulation method, multiple time step molecular dynamics (MTS-MD) of biomolecules is steered with mean solvation forces obtained from the statistical-mechanical, 3D-RISM-KH molecular theory of solvation. To speed up such hybrid simulation, mean solvation forces acting on atoms of the biomolecule in conformations at long outer time steps are obtained by converging the 3D-RISM-KH integral equations, and in conformations at successive inner time steps are calculated by generalized solvation force extrapolation (GSFE) using a relatively small number of best (closest) solute atomic coordinates and corresponding 3D-RISM-KH mean solvation forces from previous outer time steps. GSFE is based on a set of techniques including the non-Eckart like transformation of coordinate space separately for each solute atom, extension of the force-coordinate pair basis set followed by selection of the best subset, balancing the normal equations to minimize the deviations by modified least squares, and incremental increase of the outer time step in motion integration. The MTS-MD evolution of a biomolecular conformation steered with GSFE of 3D-RISM-KH mean solvation forces is efficiently stabilized with the method of optimized isokinetic Nosé-Hoover chain (OIN) thermostat.

The 3D-RISM-KH theory yields the solvation structure in terms of 3D maps of density distribution functions of solvent interaction sites around a solute molecule with full and consistent account for effects of chemical functionalities of all solution species. The solvation free energy and subsequent thermodynamics is then obtained at once as a simple integral of the correlation functions by performing thermodynamic integration analytically. The latter allows analytical differentiation of the free energy functional and thus self-consistent field coupling with MD.

Based on the first principles of statistical mechanics, 3D-RISM-KH mean solvation forces consistently reproduce, at the level of **fully converged** molecular simulation, both electrostatic forces (hydrogen bonding, other association, salt bridges, dielectric and Debye screening, ion localization) and nonpolar solvation effects (desolvation, hydrophobic hydration, hydrophobic interaction), as well as subtle interplays of these such as preferential solvation, molecular recognition and ligand binding. This is very distinct from the continuum solvation schemes such as the Poisson-Boltzmann (PB) and Generalized Born (GB) models combined with the solvent accessible surface area (SASA) empirical nonpolar terms and additional volume and dispersion integral corrections, which are parameterized for hydration free energy of biomolecules but are neither really applicable to solvation

structure effects in complex confined geometries nor transferable to solvent systems with cosolvent or electrolyte solutions at physiological concentrations.

In MD steered with 3D-RISM-KH mean solvation forces, the solvent dynamics is averaged out and the protein quasidynamics, strictly speaking, differs from its true dynamics. However, the contraction of the solvent dynamics does not affect the equilibrium conformational properties of the system. Moreover, as mean solvation forces is a statistical average of detailed solvation forces quickly varying in explicit solvent, this quasidynamics automatically filters out fast and short movements of the protein due to detailed interactions with solvent molecules but keeps the overall damping and steering effect of solvation forces. In fact, it performs essential dynamics of protein in solution (in the presence of solvent, co-solvent, counterions, and other possible components), with full effect of molecular steric forces and chemical specificities, such as desolvation barrier in hydrophobic interaction and hydrogen bonding (as distinct from implicit solvent MD missing such effects).

Generally speaking, mean solvation forces obtained by statistical-mechanical averaging depend on the protein dynamics as well as conformation, which results in mutual coupling of the protein and solvent dynamics. However, the protein dynamics, in particular, essential dynamics realizing protein functions, is usually much slower that the solvent dynamics. (Mind that solvent localization and exchange have slow rate – related to diffusion and partitioning in protein confined geometries, protein conformational transitions, and binding strength – but fast local solvent motion dynamics.) Therefore, except for particular processes where the solvent dynamics is important, its coupling with the protein is weak enough and can be neglected, resulting in a quasiequilibrium solvent description. Much in this way, 3D-RISM-KH mean solvation forces are obtained from statistically averaged correlations of quasiequilibrium solvent around the protein at successive conformational snapshots, and so the solvent dynamics is entirely decoupled from the protein conformational dynamics and is contracted in the mean solvation forces. (For comparison, in the continuum solvation models, mean solvation forces are also calculated for successive frozen conformational snapshots of the protein and so do not include coupling of the solute and solvent dynamics. However, constructed empirically, they are missing important molecular effects of solvent such as desolvation barrier and hydrogen bonding directionality, which are readily represented in the 3D-RISM-KH theory.)

For systems where solvent dynamics is important or is a target property, these questions can be addressed within the recently developed method of the generalized Langevin equation (GLE) in the 3D site formalism which, using an input of equilibrium correlation functions from 3D-RISM-KH, gives time-dependent correlation functions of both protein and solvent in a single formalism.[136,137] The GLE theory is involved computationally and also requires formulation of so-called memory kernels that adequately represent both relaxation physics and effects of chemical specificities of solvent and protein.[136] Nevertheless, it constitutes a promising approach to include the solvent dynamics and its coupling with the protein in the framework of MTS-MD steered with mean solvation forces obtained from molecular theory of solvation.

The hybrid MTS-MD/OIN/GSFE/3D-RISM-KH integrator accurately reproduces conformational properties of biomolecular systems against reference simulations using conventional MD with explicit solvent, as has been validated on the examples of hydrated alanine dipeptide, an asphaltene dimer in toluene solvent, hydrated miniprotein 1L2Y, and hydrated protein G.

The MTS-MD/OIN/GSFE/3D-RISM-KH quasidynamics of hydrated protein G showed 10-fold acceleration in terms of a "direct" comparison of the number/time of inner time steps of protein evolution to that in explicit solvent MD. Furthermore, quasidynamics steered with 3D-RISM-KH mean solvation forces provides 5- to 100-fold time scale compression of protein conformational changes coupled with solvent exchange, thus achieving a huge effective speedup of protein conformational sampling by a factor of 50 to 1000 times compared to conventional MD with explicit solvent or real time dynamics.

This enabled for the first time to quickly fold the miniprotein from a fully denatured, extended state in about 60 nanoseconds of this effective quasidynamics steered with 3D-RISM-KH mean solvation forces, compared to the average physical folding time of 4 to 9 microseconds observed in experiment.

The intrinsic acceleration inherent in this quasidynamics is expected to further increase with the complexity of proteins since 3D-RISM-KH mean solvation forces properly account for effects of chemical specificities of both solute and solvent molecules in protein confinement. The hybrid MTS-MD/OIN/GSFE/3D-RISM-KH integrator can be applied to various biomolecular systems and biomaterials – such as large proteins, DNA strands, and cellulose nanocrystals – in different solvents, solvent mixtures, and electrolyte solutions.

Acknowledgments

This work was supported by the ArboraNano – the Canadian Forest NanoProducts Network, by the University of Alberta, and by the National Institute for Nanotechnology. The high performance computing resources were provided by the WestGrid – Compute/Calcul Canada national advanced computing platform. The figures were made using the VMD (Visual Molecular Dynamics) package.[138] A.K. thanks Godehard Sutmann, Johannes Grotendorst, Gerhard Gompper, and Dominik Marx – the Organizers of the Winter Juelich School in Computational Sciences 2015 – for the opportunity of presenting this lecture.

References

1. R. P. Feynman, *There's Plenty of Room at the Bottom*, Am. Phys. Soc. Meeting, Caltech, Pasadena, CA, 1959.
2. R. P. Feynman, *Tiny Machines*, Esalen Institute, Big Sur, CA, 1984.
3. M. Lundstrom, P. Cummings, and M. Alam, "Enabling and investigative tools: Theory, modeling, and simulation", in: Nanotechnology Research Directions for Societal Needs in 2020: Retrospective and Outlook, M. C. Roco, C. A. Mirkin, and M. C. Hersam, (Eds.), chapter 1, pp. 23–69.
4. J.-P. Hansen and I. McDonald, *Theory of Simple Liquids*, Elsevier Academic Press, London Burlington, MA, 3rd edition, 2006.
5. P. L. Freddolino, C. B. Harrison, Yanxin Liu, and Klaus Schulten, *Challenges in protein-folding simulations*, Nat. Phys., **6**, no. 10, 751–758, 2010.
6. B. J. Alder and T. E. Wainwright, *Studies in Molecular Dynamics. I. General Method*, J. Chem. Phys., **31**, no. 2, 459, 1959.
7. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press Oxford University Press, Oxford England New York, 1989.

8. D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd edition.
9. B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics*, vol. 14, Cambridge University Press, 2005.
10. J. A. McCammon, B. R. Gelin, and M. Karplus, *Dynamics of folded proteins*, Nature, **267**, no. 5612, 585–590, 1977.
11. C. L. Brooks, M. Karplus, and B. M. Pettitt, *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*, Adv. Chem. Phys., **71**, 1–6, 1988.
12. A. Rojnuckarin, S. Kim, and S. Subramaniam, *Brownian dynamics simulations of protein folding: Access to milliseconds time scale and beyond*, Proc. Natl. Acad. Sci. USA, **95**, no. 8, 4288–4292, 1998.
13. Y. Duan, *Pathways to a Protein Folding Intermediate Observed in a 1-Microsecond Simulation in Aqueous Solution*, Science, **282**, no. 5389, 740–744, 1998.
14. G. Hernandez, F. E. Jenney, M. W. W. Adams, and D. M. LeMaster, *Millisecond time scale conformational flexibility in a hyperthermophile protein at ambient temperature*, Proc. Natl. Acad. Sci. USA, **97**, no. 7, 3166–3170, 2000.
15. M. Karplus and J. A. McCammon, *Molecular dynamics simulations of biomolecules*, Nat. Struct. Biol., **9**, no. 9, 646–652, 2002.
16. Y. Zhang, M. H. Peters, and Y. Li, *Nonequilibrium, multiple-timescale simulations of ligand-receptor interactions in structured protein systems*, Proteins, **52**, no. 3, 339–348, 2003.
17. S. A. Adcock and J. A. McCammon, *Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins*, Chem. Rev., **106**, no. 5, 1589–1615, 2006.
18. P. L. Freddolino, F. Liu, M. Gruebele, and K. Schulten, *Ten-Microsecond Molecular Dynamics Simulation of a Fast-Folding WW Domain*, Biophys. J., **94**, no. 10, L75–L77, 2008.
19. J. L. Klepeis, K. Lindorff-Larsen, R. O. Dror, and D. E. Shaw, *Long-timescale molecular dynamics simulations of protein structure and function*, Curr. Opin. Struct. Biol., **19**, no. 2, 120–127, 2009.
20. R. F. Service, *Custom-Built Supercomputer Brings Protein Folding Into View*, Science, **330**, no. 6002, 308–309, 2010.
21. D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, and W. Wriggers, *Atomic-Level Characterization of the Structural Dynamics of Proteins*, Science, **330**, no. 6002, 341–346, 2010.
22. M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, Oxford University Press, New York, 2010.
23. S. Genheden and U. Ryde, *Will molecular dynamics simulations of proteins ever reach equilibrium?*, Phys. Chem. Chem. Phys., **14**, no. 24, 8662, 2012.
24. D. Chandler, J. D. McCoy, and S. J. Singer, *Density functional theory of nonuniform polyatomic systems. I. General formulation*, J. Chem. Phys., **85**, no. 10, 5971–5976, 1986.
25. D. Chandler, J. D. McCoy, and S. J. Singer, *Density functional theory of nonuniform polyatomic systems. II. Rational closures for integral equations*, J. Chem. Phys., **85**, no. 10, 5977–598, 1986.
26. D. Beglov and B. Roux, *An Integral Equation To Describe the Solvation of Polar Molecules in Liquid Water*, J. Phys. Chem. B, **101**, no. 39, 7821–7826, 1997.

27. A. Kovalenko and F. Hirata, *Three-dimensional density profiles of water in contact with a solute of arbitrary shape: a RISM approach*, Chem. Phys. Lett., **290**, no. 1-3, 237–244, 1998.
28. A. Kovalenko and F. Hirata, *Self-consistent description of a metal–water interface by the Kohn–Sham density functional theory and the three-dimensional reference interaction site model*, J. Chem. Phys., **110**, no. 20, 10095–10112, 1999.
29. A. Kovalenko and F. Hirata, *Potentials of mean force of simple ions in ambient aqueous solution. I. Three-dimensional reference interaction site model approach*, J. Chem. Phys., **112**, no. 23, 10391–10402, 2000.
30. A. Kovalenko and F. Hirata, *Potentials of mean force of simple ions in ambient aqueous solution. II. Solvation structure from the three-dimensional reference interaction site model approach, and comparison with simulations*, J. Chem. Phys., **112**, no. 23, 10403–10417, 2000.
31. A. Kovalenko, "Three-dimensional rism theory for molecular liquids and solid-liquid interfaces", in: Molecular Theory of Solvation, F. Hirata, (Ed.), vol. 24 of *Understanding Chemical Reactivity*, chapter 4, pp. 169–275. Kluwer Academic Publishers, Norwell, MA, USA, 2003.
32. S. Gusarov, B. S. Pujari, and A. Kovalenko, *Efficient treatment of solvation shells in 3D molecular theory of solvation*, J. Comput. Chem., **33**, no. 17, 1478–1494, 2012.
33. A. Kovalenko, *Multiscale modeling of solvation in chemical and biological nanosystems and in nanoporous materials*, Pure Appl. Chem., **85**, no. 1, 2013.
34. T. Miyata and F. Hirata, *Combination of molecular dynamics method and 3D-RISM theory for conformational sampling of large flexible molecules in solution*, J. Comput. Chem., **29**, no. 6, 871–882, 2008.
35. T. Luchko, S. Gusarov, D. R. Roe, C. Simmerling, D. A. Case, J. Tuszynski, and A. Kovalenko, *Three-Dimensional Molecular Theory of Solvation Coupled with Molecular Dynamics in Amber*, J. Chem. Theory Comput., **6**, no. 3, 607–624, 2010.
36. I. Omelyan and A. Kovalenko, *Generalised canonical–isokinetic ensemble: speeding up multiscale molecular dynamics and coupling with 3D molecular theory of solvation*, Mol. Simul., **39**, no. 1, 25–48, 2013.
37. I. Omelyan and A. Kovalenko, *Multiple time step molecular dynamics in the optimized isokinetic ensemble steered with the molecular theory of solvation: Accelerating with advanced extrapolation of effective solvation forces*, J. Chem. Phys., **139**, no. 24, 244106, 2013.
38. I. P. Omelyan and A. Kovalenko, *MTS-MD of biomolecules steered with 3D-RISM-KH mean solvation forces accelerated with generalized solvation force extrapolation*, J. Chem. Theory Comput., 2015, (accepted).
39. J. M. Antosiewicz and D. Shugar, *Poisson–Boltzmann continuum-solvation models: applications to pH-dependent properties of biomolecules*, Mol. Biosyst., **7**, no. 11, 2923, 2011.
40. W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson, *Semianalytical treatment of solvation for molecular mechanics and dynamics*, J. Am. Chem. Soc., **112**, no. 16, 6127–6129, 1990.
41. A. Onufriev, D. Bashford, and D. A. Case, *Exploring protein native states and large-scale conformational changes with a modified generalized born model*, Proteins, **55**, no. 2, 383–394, 2004.

42. A. Onufriev, "Continuum electrostatics solvent modeling with the generalized born model", in: Modeling Solvent Environments, pp. 127–165. Wiley-Blackwell, 2010.
43. J. Mongan, C. Simmerling, J. A. McCammon, D. A. Case, and A. Onufriev, *Generalized Born Model with a Simple, Robust Molecular Volume Correction*, J. Chem. Theory Comput., **3**, no. 1, 156–169, 2007.
44. J. A. Wagoner and N. A. Baker, *Assessing implicit models for nonpolar mean solvation forces: The importance of dispersion and volume terms*, Proc. Natl. Acad. Sci. USA, **103**, no. 22, 8331–8336, 2006.
45. N. Yoshida, T. Imai, S. Phongphanphanee, A. Kovalenko, and F. Hirata, *Molecular Recognition in Biomolecules Studied by Statistical-Mechanical Integral-Equation Theory of Liquids*, J. Phys. Chem. B, **113**, no. 4, 873–886, 2009.
46. T. Imai, K. Oda, A. Kovalenko, F. Hirata, and A. Kidera, *Ligand Mapping on Protein Surfaces by the 3D-RISM Theory: Toward Computational Fragment-Based Drug Design*, J. Am. Chem. Soc., **131**, no. 34, 12430–12440, 2009.
47. T. Imai, N. Miyashita, Y. Sugita, A. Kovalenko, F. Hirata, and A. Kidera, *Functionality Mapping on Internal Surfaces of Multidrug Transporter AcrB Based on Molecular Theory of Solvation: Implications for Drug Efflux Pathway*, J. Phys. Chem. B, **115**, no. 25, 8288–8295, 2011.
48. D. Nikolić, N. Blinov, D. Wishart, and A. Kovalenko, *3D-RISM-D ock : A New Fragment-Based Drug Design Protocol*, J. Chem. Theory Comput., **8**, no. 9, 3356–3372, 2012.
49. M. Tuckerman, B. J. Berne, and G. J. Martyna, *Reversible multiple time scale molecular dynamics*, J. Chem. Phys., **97**, no. 3, 1990, 1992.
50. S. J. Stuart, R. Zhou, and B. J. Berne, *Molecular dynamics with multiple time scales: The selection of efficient reference system propagators*, J. Chem. Phys., **105**, no. 4, 1426, 1996.
51. A. Kopf, W. Paul, and B. Dünweg, *Multiple time step integrators and momentum conservation*, Comput. Phys. Commun., **101**, no. 1-2, 1–8, 1997.
52. T. Schlick, E. Barth, and M. Mandziuk, *BIOMOLECULAR DYNAMICS AT LONG TIMESTEPS:Bridging the Timescale Gap Between Simulation and Experimentation*, Annu. Rev. Biophys. Biomol. Str., **26**, no. 1, 181–222, 1997.
53. M. Watanabe and M. Karplus, *Simulations of Macromolecules by Multiple Time-Step Methods*, J. Phys. Chem., **99**, no. 15, 5680–5697, 1995.
54. M. Mandziuk and T. Schlick, *Resonance in the dynamics of chemical systems simulated by the implicit midpoint scheme*, Chem. Phys. Lett., **237**, no. 5-6, 525–535, 1995.
55. E. Barth and T. Schlick, *Extrapolation versus impulse in multiple-timestepping schemes. II. Linear analysis and applications to Newtonian and Langevin dynamics*, J. Chem. Phys., **109**, no. 5, 1633, 1998.
56. T. Schlick, M. Mandziuk, R. D. Skeel, and K. Srinivas, *Nonlinear Resonance Artifacts in Molecular Dynamics Simulations*, J. Comput. Phys., **140**, no. 1, 1–29, 1998.
57. Q. Ma, J. A. Izaguirre, and R. D. Skeel, *Verlet-I/R-RESPA/Impulse is Limited by Nonlinear Instabilities*, SIAM J. Sci. Comput., **24**, no. 6, 1951–1973, 2003.
58. I. P. Omelyan, *Processed splitting algorithms for rigid-body molecular dynamics simulations*, Phys. Rev. E, **78**, 026702–6, 2008.
59. I. P. Omelyan, *Advanced multiple time scale molecular dynamics*, J. Chem. Phys., **131**, no. 10, 104101, 2009.

60. I. P. Omelyan and A. Kovalenko, *Multiple time scale molecular dynamics for fluids with orientational degrees of freedom. I. Microcanonical ensemble*, J. Chem. Phys., **135**, no. 11, 114110–10, 2011.
61. R. J. Loncharich, B. R. Brooks, and R. W. Pastor, *Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanyl-N'-methylamide*, Biopolymers, **32**, no. 5, 523–535, 1992.
62. E. Barth and T. Schlick, *Overcoming stability limitations in biomolecular dynamics. I. Combining force splitting via extrapolation with Langevin dynamics in LN*, J. Chem. Phys., **109**, no. 5, 1617, 1998.
63. P. Minary, M. Tuckerman, and G. Martyna, *Long Time Molecular Dynamics for Enhanced Conformational Sampling in Biomolecular Systems*, Phys. Rev. Lett., **93**, no. 15, 2004.
64. J. B. Abrams, M. E. Tuckerman, and G. J. Martyna, "Equilibrium statistical mechanics, non-hamiltonian molecular dynamics, and novel applications from resonance-free timesteps to adiabatic free energy dynamics", in: Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology, pp. 139–192. 2006.
65. P. Minary, G. J. Martyna, and M. E. Tuckerman, *Algorithms and novel applications based on the isokinetic ensemble. I. Biophysical and path integral molecular dynamics*, J. Chem. Phys., **118**, no. 6, 2510, 2003.
66. I. P. Omelyan and A. Kovalenko, *Multiple time scale molecular dynamics for fluids with orientational degrees of freedom. II. Canonical and isokinetic ensembles*, J. Chem. Phys., **135**, no. 23, 234107, 2011.
67. I. P. Omelyan and A. Kovalenko, *Overcoming the Barrier on Time Step Size in Multiscale Molecular Dynamics Simulation of Molecular Liquids*, J. Chem. Theory Comput., **8**, no. 1, 6–16, 2012.
68. J. G. Kirkwood, *Statistical Mechanics of Fluid Mixtures*, J. Chem. Phys., **3**, no. 5, 300, 1935.
69. Donald A McQuarrie, *Statistical Mechanics*, University Science Books, Sausalito, C., 2000.
70. M. C. Stumpe, N. Blinov, D. Wishart, A. Kovalenko, and V. S. Pande, *Calculation of Local Water Densities in Biological Systems: A Comparison of Molecular Dynamics Simulations and the 3D-RISM-KH Molecular Theory of Solvation*, J. Phys. Chem. B, **115**, no. 2, 319–328, 2011.
71. S Gusarov, T. Ziegler, and A. Kovalenko, *Self-Consistent Combination of the Three-Dimensional RISM Theory of Molecular Solvation with Analytical Gradients and the Amsterdam Density Functional Package*, J. Phys. Chem. A, **110**, no. 18, 6083–6090, 2006.
72. D. Casanova, S. Gusarov, A. Kovalenko, and T. Ziegler, *Evaluation of the SCF Combination of KS-DFT and 3D-RISM-KH Solvation Effect on Conformational Equilibria, Tautomerization Energies, and Activation Barriers*, J. Chem. Theory Comput., **3**, no. 2, 458–476, 2007.
73. J. W. Kaminski, S. Gusarov, T. A. Wesolowski, and A. Kovalenko, *Modeling Solvatochromic Shifts Using the Orbital-Free Embedding Potential at Statistically Mechanically Averaged Solvent Density*, J. Phys. Chem. A, **114**, no. 20, 6082–6096, 2010.
74. T. Yamazaki and A. Kovalenko, *Spatial Decomposition Analysis of the Thermodynamics of Cyclodextrin Complexation*, J. Chem. Theory Comput., **5**, no. 7, 1723–1730, 2009.

75. L. M. da Costa, S. Hayaki, S. R. Stoyanov, S. Gusarov, X. Tan, M. R. Gray, J. M. Stryker, R. Tykwinski, J. W. de M. Carneiro, H. Sato, P. R. Seidl, and A. Kovalenko, *3D-RISM-KH molecular theory of solvation and density functional theory investigation of the role of water in the aggregation of model asphaltenes*, Phys. Chem. Chem. Phys., **14**, no. 11, 3922, 2012.
76. S. R. Stoyanov, S. Gusarov, and A. Kovalenko, "Multiscale modeling of the adsorption interaction between model bitumen compounds and zeolite nanoparticles in gas and liquid phase.", in: Industrial applications of molecular simulations, Marc Meunier, (Ed.), chapter 14. CRC Press, Boca Raton, FL, 2012.
77. J. Fafard, O. Lyubimova, S. R. Stoyanov, G. K. Dedzo, S. Gusarov, A. Kovalenko, and C. Detellier, *Adsorption of Indole on Kaolinite in Nonaqueous Media: Organoclay Preparation and Characterization, and 3D-RISM-KH Molecular Theory of Solvation Investigation*, J. Phys. Chem. C, **117**, no. 36, 18556–18566, 2013.
78. W.-J. Huang, G. K. Dedzo, S. R. Stoyanov, O. Lyubimova, S. Gusarov, S. Singh, H. Lao, A. Kovalenko, and C. Detellier, *Molecule–Surface Recognition between Heterocyclic Aromatic Compounds and Kaolinite in Toluene Investigated by Molecular Theory of Solvation and Thermodynamic and Kinetic Experiments*, J. Phys. Chem. C, **118**, no. 41, 23821–23834, 2014.
79. A. Kovalenko, A. E. Kobryn, S. Gusarov, O. Lyubimova, X.-J. Liu, N. Blinov, and M. Yoshida, *Molecular theory of solvation for supramolecules and soft matter structures: application to ligand binding, ion channels, and oligomeric polyelectrolyte gelators*, Soft Matter, **8**, no. 5, 1508–1520, 2012.
80. J. G. Moralez, J. Raez, T. Yamazaki, R. K. Motkuri, A. Kovalenko, and H. Fenniri, *Helical Rosette Nanotubes with Tunable Stability and Hierarchy*, J. Am. Chem. Soc., **127**, no. 23, 8307–8309, 2005.
81. R. S. Johnson, T. Yamazaki, A. Kovalenko, and H. Fenniri, *Molecular Basis for Water-Promoted Supramolecular Chirality Inversion in Helical Rosette Nanotubes*, J. Am. Chem. Soc., **129**, no. 17, 5735–5743, 2007.
82. T. Yamazaki, H. Fenniri, and A. Kovalenko, *Structural Water Drives Self-assembly of Organic Rosette Nanotubes and Holds Host Atoms in the Channel*, ChemPhysChem, **11**, no. 2, 361–367, 2010.
83. R. L. Silveira, S. R. Stoyanov, S. Gusarov, M. S. Skaf, and A. Kovalenko, *Plant Biomass Recalcitrance: Effect of Hemicellulose Composition on Nanoscale Forces that Control Cell Wall Strength*, J. Am. Chem. Soc., **135**, no. 51, 19048–19051, 2013.
84. A. Kovalenko, *Computational modeling of the structure relaxation and dispersion thermodynamics of pristine and modified cellulose nanocrystals in solution*, Nord. Pulp Pap. Res. J., **29**, no. 01, 144–155, 2014.
85. R. L. Silveira, S. R. Stoyanov, S. Gusarov, M. S. Skaf, and A. Kovalenko, *Supramolecular Interactions in Secondary Plant Cell Walls: Effect of Lignin Chemical Composition Revealed with the Molecular Theory of Solvation*, J. Phys. Chem. Lett., **6**, no. 1, 206–211, 2015.
86. S. Phongphanphanee, N. Yoshida, and F. Hirata, *On the Proton Exclusion of Aquaporins: A Statistical Mechanics Study*, J. Am. Chem. Soc., **130**, no. 5, 1540–1541, 2008.
87. Y. Kiyota, R. Hiraoka, N. Yoshida, Y. Maruyama, T. Imai, and F. Hirata, *Theoretical Study of CO Escaping Pathway in Myoglobin with the 3D-RISM Theory*, J. Am. Chem. Soc., **131**, no. 11, 3852–3853, 2009.

88. Q. Li, S. Gusarov, S. Evoy, and A. Kovalenko, *Electronic Structure, Binding Energy, and Solvation Structure of the Streptavidin-Biotin Supramolecular Complex: ONIOM and 3D-RISM Study*, J. Phys. Chem. B, **113**, no. 29, 9958–9967, 2009.
89. S. Genheden, T. Luchko, S. Gusarov, A. Kovalenko, and U. Ryde, *An MM/3D-RISM Approach for Ligand Binding Affinities*, J. Phys. Chem. B, **114**, no. 25, 8505–8516, 2010.
90. N. Blinov, L. Dorosh, D. Wishart, and A. Kovalenko, *Association Thermodynamics and Conformational Stability of β-Sheet Amyloid β(17-42) Oligomers: Effects of E22Q (Dutch) Mutation and Charge Neutralization*, Biophys. J., **98**, no. 2, 282–296, 2010.
91. Y. Maruyama, N. Yoshida, and F. Hirata, *Revisiting the Salt-Induced Conformational Change of DNA with 3D-RISM Theory*, J. Phys. Chem. B, **114**, no. 19, 6464–6471, 2010.
92. S. Phongphanphanee, T. Rungrotmongkol, N. Yoshida, S. Hannongbua, and F. Hirata, *Proton Transport through the Influenza A M2 Channel: Three-Dimensional Reference Interaction Site Model Study*, J. Am. Chem. Soc., **132**, no. 28, 9782–9788, 2010.
93. A. Kovalenko and N. Blinov, *Multiscale methods for nanochemistry and biophysics in solution*, J. Mol. Liq., **164**, no. 1-2, 101–112, 2011.
94. T. Yamazaki and A. Kovalenko, *Spatial Decomposition of Solvation Free Energy Based on the 3D Integral Equation Theory of Molecular Liquid: Application to Miniproteins*, J. Phys. Chem. B, **115**, no. 2, 310–318, 2011.
95. T. Yamazaki and H. Fenniri, *Imaging Carbon Nanotube Interaction with Nucleobases in Water Using the Statistical Mechanical Theory of Molecular Liquids*, J. Phys. Chem. C, **116**, no. 28, 15087–15092, 2012.
96. A. Kovalenko, "Partial molar volumes of proteins in solution: Prediction by statistical-mechanical, 3d-rism-kb molecular theory of solvation", chapter 22, pp. 575–610. Royal Society of Chemistry, 2015.
97. J. S. Perkyns and B. M. Pettitt, *A dielectrically consistent interaction site theory for solvent—electrolyte mixtures*, Chem. Phys. Lett., **190**, no. 6, 626–630, 1992.
98. J. S. Perkyns and B. M. Pettitt, *A site–site theory for finite concentration saline solutions*, J. Chem. Phys., **97**, no. 10, 7656, 1992.
99. A. Kovalenko and F. Hirata, *First-principles realization of a van der Waals–Maxwell theory for water*, Chem. Phys. Lett., **349**, no. 5-6, 496–502, 2001.
100. A. Kovalenko and F. Hirata, *Towards a Moleculat Theory for the van der Waals Maxwell Description of Fluid Phase Transitions*, J. Theor. Comput. Chem., **01**, no. 02, 381–406, 2002.
101. K. Yoshida, T. Yamaguchi, A. Kovalenko, and F. Hirata, *Structure of tert -Butyl Alcohol-Water Mixtures Studied by the RISM Theory*, J. Phys. Chem. B, **106**, no. 19, 5042–5049, 2002.
102. T. Yamazaki, A. Kovalenko, V. V. Murashov, and G. N. Patey, *Ion Solvation in a Water-Urea Mixture*, J. Phys. Chem. B, **114**, no. 1, 613–619, 2010.
103. A. E. Kobryn, D. Nikolić, O. Lyubimova, S. Gusarov, and A. Kovalenko, *Dissipative Particle Dynamics with an Effective Pair Potential from Integral Equation Theory of Molecular Liquids*, J. Phys. Chem. B, **118**, no. 41, 12034–12049, 2014.
104. G. Schmeer and A. Maurer, *Development of thermodynamic properties of electrolyte solutions with the help of RISM-calculations at the Born–Oppenheimer level*, Phys. Chem. Chem. Phys., **12**, no. 10, 2407, 2010.

105. C. Chipot and A. Pohorille, *Free energy calculations: theory and applications in chemistry and biology*, 2007.
106. J. S. Perkyns, G. C. Lynch, J. J. Howard, and B. M. Pettitt, *Protein solvation from theory and simulation: Exact treatment of Coulomb interactions in three-dimensional theories*, J. Chem. Phys., **132**, no. 6, 064106, 2010.
107. A. Kovalenko, S. Ten-no, and F. Hirata, *Solution of three-dimensional reference interaction site model and hypernetted chain equations for simple point charge water by modified method of direct inversion in iterative subspace*, J. Comput. Chem., **20**, no. 9, 928–936, 1999.
108. J.-F. Truchon, B. M. Pettitt, and P. Labute, *A Cavity Corrected 3D-RISM Functional for Accurate Solvation Free Energies*, J. Chem. Theory Comput., **10**, no. 3, 934–941, 2014.
109. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, *A smooth particle mesh Ewald method*, J. Chem. Phys., **103**, no. 19, 8577, 1995.
110. I. P. Omelyan, *Ewald summation technique for interaction site models of polar fluids*, Comput. Phys. Commun., **107**, no. 1-3, 113–122, 1997.
111. I. P. Omelyan, *A New Leapfrog Integrator of Rotational Motion. The Revised Angular-Momentum Approach*, Mol. Simul., **22**, no. 3, 213–236, 1999.
112. C. Lawson, *Solving Least Squares Problems*, SIAM, Philadelphia, 1995.
113. G. Quintana-Ortí, E. S. Quintana-Ortí, and A. Petitet, *Efficient Solution Of The Rank-Deficient Linear Least Squares Problem*, SIAM J. Sci. Comput., **20**, no. 3, 1155–1163, 1998.
114. G. R. Kneller, *Eckart axis conditions, Gauss' principle of least constraint, and the optimal superposition of molecular structures*, J. Chem. Phys., **128**, no. 19, 194101, 2008.
115. C. Eckart, *Some Studies Concerning Rotating Axes and Polyatomic Molecules*, Phys. Rev., **47**, no. 7, 552–558, 1935.
116. J. D. Louck and H. W. Galbraith, *Eckart vectors, Eckart frames, and polyatomic molecules*, Rev. Mod. Phys., **48**, 69–106, 1976.
117. D. Janežič, M. Praprotnik, and F. Merzel, *Molecular dynamics integration and molecular vibrational theory. I. New symplectic integrators*, J. Chem. Phys., **122**, no. 17, 174101, 2005.
118. M. Praprotnik and D. Janežič, *Molecular dynamics integration and molecular vibrational theory. II. Simulation of nonlinear molecules*, J. Chem. Phys., **122**, no. 17, 174102, 2005.
119. M. Praprotnik and D. Janežič, *Molecular dynamics integration and molecular vibrational theory. III. The infrared spectrum of water*, J. Chem. Phys., **122**, no. 17, 174103, 2005.
120. I. Omelyan and A. Kovalenko, *Interpretation of atomic motion in flexible molecules: Accelerating molecular dynamics simulations*, Phys. Rev. E, **85**, no. 2, 2012.
121. E. A. Coutsias, C. Seok, and K. A. Dill, *Using quaternions to calculate RMSD*, J. Comput. Chem., **25**, no. 15, 1849–1857, 2004.
122. P. Liu, D. K. Agrafiotis, and D. L. Theobald, *Fast determination of the optimal rotational matrix for macromolecular superpositions*, J. Comput. Chem., pp. 1561–1563, 2009.
123. G. Chevrot, P. Calligari, K. Hinsen, and G. R. Kneller, *Least constraint approach to the extraction of internal motions from molecular dynamics trajectories of flexible macromolecules*, J. Chem. Phys., **135**, no. 8, 084110, 2011.

124. D. A. Case, T. A. Darden, III T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. P. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P. A. Kollman, "Amber 11", University of California, San Francisco, 2010.
125. I. P. Omelyan, I. M. Mryglod, and R. Folk, *Optimized Verlet-like algorithms for molecular dynamics simulations*, Phys. Rev. E, **65**, 056706–5, 2002.
126. J. W. Neidigh, R. M. Fesinmeyer, and N. H. Andersen, *Designing a 20-residue protein*, Nat. Struct. Biol., **9**, no. 6, 425–430, 2002.
127. T. S. Ulmer, B. E. Ramirez, F. Delaglio, and A. Bax, *Evaluation of Backbone Proton Positions and Dynamics in a Small Protein by Liquid Crystal NMR Spectroscopy*, J. Am. Chem. Soc., **125**, no. 30, 9179–9191, 2003.
128. J. W. Pitera and W. Swope, *Understanding folding and design: Replica-exchange simulations of "Trp-cage" miniproteins*, Proc. Natl. Acad. Sci. USA, **100**, no. 13, 7587–7592, 2003.
129. D. Paschek, H. Nymeyer, and A. E. García, *Replica exchange simulation of reversible folding/unfolding of the Trp-cage miniprotein in explicit solvent: On the structure and possible role of internal water*, J. Struct. Biol., **157**, no. 3, 524–533, 2007.
130. R. Day, D. Paschek, and A. E. Garcia, *Microsecond simulations of the folding/unfolding thermodynamics of the Trp-cage miniprotein*, Proteins, pp. 1889—-1899, 2010.
131. A. S. N. Seshasayee, Theor. Biol. Med. Model., **2**, no. 1, 7, 2005.
132. C. Simmerling, B. Strockbine, and A. E. Roitberg, *All-Atom Structure Prediction and Folding Simulations of a Stable Protein*, J. Am. Chem. Soc., **124**, no. 38, 11258–11259, 2002.
133. L. Qiu, S. A. Pabit, A. E. Roitberg, and S. J. Hagen, *Smaller and Faster: The 20-Residue Trp-Cage Protein Folds in 4 μs*, J. Am. Chem. Soc., **124**, no. 44, 12952–12953, 2002.
134. C. D. Snow, B. Zagrovic, and V. S. Pande, *The Trp Cage: Folding Kinetics and Unfolded State Topology via Molecular Dynamics Simulations*, J. Am. Chem. Soc., **124**, no. 49, 14548–14549, 2002.
135. Q. Shao, J. Shi, and W. Zhu, *Enhanced sampling molecular dynamics simulation captures experimentally suggested intermediate and unfolded states in the folding pathway of Trp-cage miniprotein*, J. Chem. Phys., **137**, no. 12, 125103, 2012.
136. B. Kim and F. Hirata, *Structural fluctuation of protein in water around its native state: A new statistical mechanics formulation*, J. Chem. Phys., **138**, no. 5, 054108, 2013.
137. F. Hirata and K. Akasaka, *Structural fluctuation of proteins induced by thermodynamic perturbation*, J. Chem. Phys., **142**, no. 4, 044110, 2015.
138. W. Humphrey, A. Dalke, and K. Schulten, *VMD – Visual Molecular Dynamics*, J. Mol. Graphics, **14**, 33–38, 1996.

Integral Equation Theory as a Solvation Model for Classical and Quantum Solute Systems

Stefan M. Kast, Jochen Heil, Franziska Hoffgaard

Physikalische Chemie III, Fakultät für Chemie und Chemische Biologie
Technische Universität Dortmund, 44227 Dortmund, Germany

E-mail: {*stefan.kast, jochen.heil, franziska.hoffgaard*}*@tu-dortmund.de*

We review the capabilities of liquid state theory as a solvation model for solute systems described by various levels of theory. Starting with a characterization of the pure solvent in terms of its site-site intermolecular distribution functions, the chemical potential of the solute can be computed analytically by solving a solute-solvent integral equation theory for a given solute-solvent Hamiltonian. The effect of electronic polarizability can be captured by allowing either the electronic wave function or a classical model of intramolecular polarizability to adapt to a polarizing environment self-consistently. The target quantity, the free energy in solution is then given by adding the relaxed intramolecular energy to the chemical potential originating from the correspondingly adapted intermolecular interactions. We outline integral equation theory in general and specialize on realizations of such schemes based on the three-dimensional "reference interaction site model" (3D RISM), supplemented by illustrations of the performance for a range of benchmark systems.

1 Introduction

1.1 Solvation Effects

Solvation affects the thermodynamics as well as the dynamics of a system that is immersed in a fluid environment. Physically, the presence of a solvent alters the effective interactions between and within solute species, giving rise to a "free energy surface" that is defined by adding the intramolecular energy and the chemical potential. The latter is essentially a measure for the free energy change upon moving the solute from the (ideal) gas phase into the solvent phase (where we have to clearly specify reference conditions and the molecular state). The driving force of molecular processes involving the solute is therefore not only its intramolecular Hamiltonian but in addition the change in free energy of the *total* system including the solvent. Dynamically, if solvent degrees of freedom are supposed to relax faster than intra-solute modes, one would need to modify equations of motion accordingly, for instance by accounting for effective solvent friction and random forces in the sense of the Langevin equation. In these notes, we restrict ourselves to the thermodynamics.

1.2 Statistical Mechanics of Solvation

We start with the basic statistical-mechanical reduction of a system composed of solute (u) and solvent (v) particles and their respective coordinates $\mathbf{R}_u$ and $\mathbf{R}_v$ (for further details see for instance[1,2,3]). In thermal equilibrium, working in the canonical ensemble, the expression

$$P(\mathbf{R}_u, \mathbf{R}_v) = \frac{\exp[-\beta U(\mathbf{R}_u, \mathbf{R}_v)]}{\int \exp[-\beta U(\mathbf{R}_u, \mathbf{R}_v)]\, \mathrm{d}\mathbf{R}_u\, \mathrm{d}\mathbf{R}_v} \tag{1}$$

represents the probability of finding a specific solute-solvent configuration. It is given by the normalized Boltzmann factor of the configuration's total potential energy U scaled by an inverse temperature β. The key to a reduction to an effective energy function is based on the formal procedure of averaging (integrating) out solvent degrees of freedom via

$$P(\mathbf{R}_u) = \frac{\exp[-\beta W(\mathbf{R}_u)]}{\int \exp[-\beta W(\mathbf{R}_u)]\,\mathrm{d}\mathbf{R}_u} \tag{2}$$

where we define the "potential of mean force" (PMF) W by

$$\exp[-\beta W(\mathbf{R}_u)] = \int \exp[-\beta U(\mathbf{R}_u, \mathbf{R}_v)]\,\mathrm{d}\mathbf{R}_v. \tag{3}$$

The PMF constitutes the *effective* potential energy controlling the solute structure (where the term "solute" can refer to an arbitrary collection of diluted species such that for instance association reactions can be treated on the same footings as intramolecular processes). It is clearly a *free energy*, since it can be shown that the PMF is related to an ensemble average of a derivative of the potential energy by the "reversible work theorem"[4],

$$\frac{\partial W(\mathbf{R}_u)}{\partial \mathbf{R}_{u,i}} = \left\langle \frac{\partial U(\mathbf{R}_u, \mathbf{R}_v)}{\partial \mathbf{R}_{u,i}} \right\rangle \equiv -\left\langle \mathbf{f}_i \right\rangle, \tag{4}$$

with the "mean force" $\mathbf{f}$ acting on a specific solute degree of freedom i. In this sense, the difference of the Helmholtz (free) energy F between two (frozen) solute configurations in an always fully relaxed, equilibrated solvent is directly related to the difference of the PMF,

$$\begin{aligned} W(\mathbf{R}_{u(2)}) - W(\mathbf{R}_{u(1)}) &= F(\mathbf{R}_{u(2)}) - F(\mathbf{R}_{u(1)}) \\ &= \mu^{\mathrm{ex}}(\mathbf{R}_{u(2)}) + U(\mathbf{R}_{u(2)}) - [\mu^{\mathrm{ex}}(\mathbf{R}_{u(1)}) + U(\mathbf{R}_{u(1)})], \end{aligned} \tag{5}$$

which in turn is equivalent to the sum of the change of intra-solute energy and the respective excess chemical potentials μ^{ex} in the given configurations. The latter identity follows from constructing a thermodynamic cycle where the solute is moved in its frozen configuration and with formally frozen intramolecular energy (i.e. ignoring a change upon moving it into a differently polarizing environment) from the ideal gas phase into solution. Thus, for an unpolarizable intramolecular energy function, μ^{ex} equals to the solvation free energy (assuming identical reference concentrations in both phases), and the intramolecular energy in solution and in the gas phase would be the same.

By viewing the PMF as a free energy difference between two fixed (i.e. constrained as denoted by the Dirac delta function δ) configurations of a solute molecule, it is possible to connect it with distribution functions g,

$$\begin{aligned} W(\mathbf{R}_{u(2)}) - W(\mathbf{R}_{u(1)}) &= -\beta^{-1} \ln \frac{\int \delta(\mathbf{R}_u - \mathbf{R}_{u(2)}) \exp[-\beta U(\mathbf{R}_u, \mathbf{R}_v)]\,\mathrm{d}\mathbf{R}_u\,\mathrm{d}\mathbf{R}_v}{\int \delta(\mathbf{R}_u - \mathbf{R}_{u(1)}) \exp[-\beta U(\mathbf{R}_u, \mathbf{R}_v)]\,\mathrm{d}\mathbf{R}_u\,\mathrm{d}\mathbf{R}_v} \\ &= -\beta^{-1} \ln \frac{P(\mathbf{R}_{u(2)})}{P(\mathbf{R}_{u(1)})} \equiv -\beta^{-1} \ln \frac{g(\mathbf{R}_{u(2)})}{g(\mathbf{R}_{u(1)})}, \end{aligned} \tag{6}$$

which simplifies for, e.g., two atoms at a distance $|\mathbf{R}_{u(2)} - \mathbf{R}_{u(1)}| = r$ to

$$W(r) = -\beta^{-1} \ln g^{(2)}(r) \equiv g(r) \tag{7}$$

where we assumed $g^{(2)} \to 1$ for infinite separation in the thermodynamic limit and the superscript "(2)" indicates the pair distribution function (PDF). Note that the general relation between probability and distribution function involves consideration of equivalent atoms in a fluid mixture, i.e. we are free to attribute particles to types "solute" and "solvent" on chemical or topological grounds. For instance, we could single out two solvent molecules (each for simplicity assumed to be a single site here) in a pure solvent and designate them as one "solute supermolecule". By taking into account that we have N possibilities for picking the first molecule among a total of N and N-1 for the second, we have[4]

$$N(N-1)P(r) = \rho_0^2 g^{(2)}(r) \equiv \rho_0^2 g^{vv}(r) \tag{8}$$

by normalizing to the bulk solvent density $\rho_0 = N/V$ where V is the volume. In this case, the superscript "vv" denotes the solvent-solvent PDF. Analogously, for a single solute (again assumed to be single site) in a solvent of N-1 particles, we obtain (solute-solvent, uv case)

$$(N-1)\,P(r) = \frac{\rho_0}{V} g^{uv}(r), \tag{9}$$

and for two single solute species (solute-solute, uu case)

$$P(r) = \frac{1}{V^2} g^{uu}(r). \tag{10}$$

In the thermodynamic limit for $N \to \infty$ and $V \to \infty$ with ρ_0 being finite the latter two situations essentially imply zero solute density.

The preceding derivations show that access to the free energy surface in solution is provided by computing distribution functions. However, being a state function, the free energy can also be calculated according to Eq. (5) directly by adding the supermolecule's excess chemical potential (based on its – possibly polarized – solute-solvent Hamiltonian in solution) and the solute-solute interaction potential, again possibly modified due to the presence of the solvent in comparison with gas phase energies. The PDFs defined for the solute-solvent case are also linked to μ^{ex} via the coupling parameter approach that forms the basis of "thermodynamic integration" methods for computing solvation free energies[5,6]. For general anisotropic rigid bodies with relative distance vector $\mathbf{r}$ and solid angle $\mathbf{\Omega}$, the expression reads

$$\mu^{\text{ex}} = \int_0^1 \left\langle \frac{\partial U^{uv}(\lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda = \rho_0 \int_0^1 \mathrm{d}\lambda \int \mathrm{d}\mathbf{r}\, \mathrm{d}\mathbf{\Omega}\; g^{uv}(\mathbf{r}, \mathbf{\Omega}; \lambda) \frac{\partial U^{uv}(\mathbf{r}, \mathbf{\Omega}; \lambda)}{\partial \lambda} \tag{11}$$

where the solute-solvent interaction U^{uv} is assumed to vanish for $\lambda = 0$ and to be fully effective for $\lambda = 1$. A hierarchy of reductions to pairs of solute molecule and solvent sites (γ) or even a complete decomposition into all solute (α) and solvent site pairs is possible, yielding equivalently

$$\mu^{\text{ex}} = \int_0^1 \mathrm{d}\lambda \sum_\gamma \rho_{0,\gamma} \int \mathrm{d}\mathbf{r}\, g_\gamma^{uv}(\mathbf{r}; \lambda) \frac{\partial U_\gamma^{uv}(\mathbf{r}; \lambda)}{\partial \lambda} \tag{12}$$

and

$$\mu^{\text{ex}} = \int_0^1 \mathrm{d}\lambda \sum_\gamma \rho_{0,\gamma} \sum_\alpha \int \mathbf{dr}\, g^{uv}_{\alpha\gamma}(r;\lambda) \frac{\partial U^{uv}_{\alpha\gamma}(r;\lambda)}{\partial \lambda} \tag{13}$$

under the assumptions of pairwise additivity of solute-solvent interactions. These expressions show that PDFs are the key to a complete thermodynamic description of solvated systems.

2 Integral Equation Theory

2.1 Concepts

While PDFs are in principle accessible from MD (or Monte-Carlo) simulations, such sampling approaches suffer from both, statistical noise and ergodicity, i.e. limited time scale problems. Particularly for slow processes characterized by large barriers, alternative strategies can be useful. In this context, liquid state theory provides the framework to predict PDFs directly, noise-free, and analytically from interaction potentials. Given that the many-body problem cannot be solved exactly, it is inevitable that such theories have approximate character. In this sense, a sensible and practically useful theory should on one hand predict liquid and solution structure along with thermodynamics with an accuracy that approaches simulation data. On the other hand, since the time scale is not an issue, liquid state theory has the additional benefit that it can be coupled with molecular electronic structure calculations in a way that is barely possible with simulations.

Essentially two liquid state theory frameworks are practically important for molecular systems in solution, classical density functional theory (cDFT)[7,8,9] and integral equation theory which can be conceptually derived from the former[7]. In this chapter we will first outline the fundamental theory of integral equations for the simple liquid case, followed by the extension to molecular systems and quantum chemistry in the next.

2.2 Ornstein-Zernike and Closure Equations

The Ornstein-Zernike (OZ) equation connects the "total (pair) correlation function" (TCF) $h = g$–1 with the "direct correlation function" (DCF) c via a convolution product that reads for a simple homogeneous liquid composed of spherically symmetric particles 1 and 2

$$t(r = |\mathbf{r}_{12}|) = h(r) - c(r) = \rho_0 \int \mathbf{dr}_3 c(r_{32}) h(r_{13}) \equiv \rho_0 (c * h)(r) \tag{14}$$

where t represents the "indirect correlation function" (ICF). Generalization to inhomogeneous solutions[10] (for instance due to the presence of a wall or confinement) or to anisotropic molecular systems with orientational degrees of freedom (molecular Ornstein-Zernike equation, MOZ[11]) is possible but will not be covered here. This relation can be viewed as constitutive for the DCF, whose physical meaning can be interpreted as that part of the total correlation between particles not mediated by others, which becomes apparent by inserting h recursively on the r.h.s. of Eq. (14). Its range is therefore comparable to that of the interaction potential.

While we note but not elaborate here the existence of alternative derivations from diagrammatic expansions[7], the cDFT formalism provides the most modern framework for understanding the origin. Briefly, the key relation in cDFT is the stationarity of the grand potential written as functional over the (inhomogeneous) particle density, at the equilibrium density for a given (external) potential that perturbs the homogeneous liquid structure

$$\Omega = F - \mu N = F[\rho(\mathbf{r})] - \mu \int \mathrm{d}\mathbf{r}\, \rho(\mathbf{r}) \tag{15}$$

with the functional derivative with respect to the density

$$0 = \left. \frac{\delta \Omega}{\delta \rho(\mathbf{r})} \right|_{\rho=\rho(\mathrm{eq})}. \tag{16}$$

The potential can be realized, e.g., by an external field acting on the particles, but can also be viewed as originating from one of the liquid's particles. In this sense (known in the literature as "Percus' trick"), the framework originally designed for inhomogeneous systems can also be used for treating homogeneous liquids. The potential U_{ext} enters the Helmholtz free energy functional that can be decomposed into an intrinsic (subscript "int") part and the effect of the external potential. The first term is comprised of a (known) ideal and a (generally unknown) excess part, giving rise to the expression

$$\begin{aligned} F[\rho] &= F_{\mathrm{int}}[\rho] + \int \mathrm{d}\mathbf{r}\, \rho(\mathbf{r})\, U_{\mathrm{ext}}(\mathbf{r}) \\ &= F_{\mathrm{int}}^{\mathrm{ideal}}[\rho] + F_{\mathrm{int}}^{\mathrm{ex}}[\rho] + \int \mathrm{d}\mathbf{r}\, \rho(\mathbf{r})\, U_{\mathrm{ext}}(\mathbf{r}) \\ &= F_{\mathrm{int}}^{\mathrm{ideal}}[\rho] + F_{\mathrm{int}}^{\mathrm{ex,hom.ref.}}[\rho] + F_{\mathrm{int}}^{\mathrm{ex},B}[\rho] + \int \mathrm{d}\mathbf{r}\, \rho(\mathbf{r})\, U_{\mathrm{ext}}(\mathbf{r}). \end{aligned} \tag{17}$$

The stationarity of the grand potential then implies

$$\left. \frac{\delta F}{\delta \rho(\mathbf{r})} \right|_{\rho=\rho(\mathrm{eq})} - \mu + V_{\mathrm{ext}}(\mathbf{r}) = 0. \tag{18}$$

Approximations have to be made for $F_{\mathrm{int}}^{\mathrm{ex}}[\rho]$. A useful way is to split this term into a part termed "homogeneous reference" component $F_{\mathrm{int}}^{\mathrm{ex,hom.ref.}}[\rho]$and a rest $F_{\mathrm{int}}^{\mathrm{ex},B}[\rho]$ that can either be approximated further or neglected. The first part can be calculated analytically to second order in a density expansion within the conceptual framework of cDFT, which defines h and c by functional derivatives of F and Ω (where we implicitly assume that the derivatives are taken at the equilibrium density in order to obtain the equilibrium TCF and DCF):

$$\frac{\delta F_{\mathrm{int}}^{\mathrm{ex}}[\rho]}{\delta \rho(\mathbf{r})} = \mu^{\mathrm{ex}}(\mathbf{r}) \equiv -\beta^{-1} c^{(1)}(\mathbf{r}), \tag{19}$$

$$\frac{\delta^2 F_{\mathrm{int}}^{\mathrm{ex}}[\rho]}{\delta \rho(\mathbf{r})\, \delta \rho(\mathbf{r}')} = -\beta^{-1} c^{(2)}(\mathbf{r}, \mathbf{r}') \equiv -\beta^{-1} c(\mathbf{r}, \mathbf{r}'), \tag{20}$$

$$-\beta^{-1} \frac{\delta^2 \Omega[\rho]}{\delta \psi(\mathbf{r})\, \delta \psi(\mathbf{r}')} = h(\mathbf{r}, \mathbf{r}')\, \rho(\mathbf{r}) \rho(\mathbf{r}') + \rho(\mathbf{r})\, \delta(\mathbf{r} - \mathbf{r}'), \tag{21}$$

with $\psi(\mathbf{r}) = \mu(\mathbf{r}) - U_{\mathrm{ext}}(\mathbf{r})$. Tedious algebra[7,8] ultimately leads from Eqs. (18)-(21) to the OZ Eq. (14) and another so-called "closure" relation that "closes" it by providing

an independent expression connecting h and c together with the (external) potential which we now simply abbreviate by u. Assuming the validity of pairwise additive molecular interactions and the Percus trick we obtain for a simple liquid

$$h(r) + 1 = g(r) = \exp[-\beta u(r) + h(r) - c(r) + B(r)] \tag{22}$$

where the so-called "bridge function" B is defined by a functional derivative of the "remainder" free energy functional beyond second order in the density expansion according to[12]

$$B(r) = \beta \frac{\delta F_{\text{int}}^{\text{ex},B}}{\delta \rho(\mathbf{r})}\bigg|_{\rho=\rho(\text{eq})} . \tag{23}$$

Neglecting B by setting it identically to zero represents the "hypernetted chain" (HNC) approximation[13], $B_{\text{HNC}}(r) = 0$, while various non-zero approximations are available such as the Percus-Yevick (PY) closure or the mean spherical approximation (MSA)[7]. Typically and for practical purposes, as outlined further below, bridge functions are approximately written as functions (not functionals) of the ICF, as in the case of PY the closure becomes for example

$$B_{\text{PY}}(r) = \ln[1 + t(r)] - t(r), \tag{24}$$

corresponding to a linearization of the general closure exponential except for the potential.

2.3 Multicomponent Ornstein-Zernike Equations

In order to connect the OZ framework with the hierarchy of PDFs defined in chapter 1, it is useful to write down a general expression for fluid mixtures. For example, a two component mixture of spherical particles with an infinite diluted solute species in a single component solvent would yield in a compact matrix notation

$$\begin{pmatrix} h^{vv} & h^{vu} \\ h^{uv} & h^{uu} \end{pmatrix} - \begin{pmatrix} c^{vv} & c^{vu} \\ c^{uv} & c^{uu} \end{pmatrix} = \begin{pmatrix} c^{vv} & c^{vu} \\ c^{uv} & c^{uu} \end{pmatrix} * \left[\begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} h^{vv} & h^{vu} \\ h^{uv} & h^{uu} \end{pmatrix} \right] \tag{25}$$

where the order of u and v indices does not matter in this simple case. Upon performing the matrix multiplication (i.e. by defining the star between matrices as usual multiplication operator with each resulting matrix element given as a convolution), we obtain three separate equations for the pure solvent (vv), solute-solvent (uv), and solute-solute (uu) cases[2,14,15,16] (derived in these references for the more general molecular case discussed below),

$$h^{vv}(r) = c^{vv}(r) + (c^{vv} * \rho_0 h^{vv})(r), \tag{26}$$

$$h^{uv}(r) = c^{uv}(r) + (c^{uv} * \rho_0 h^{vv})(r), \tag{27}$$

$$h^{uu}(r) = c^{uu}(r) + (c^{uv} * \rho_0 h^{vu})(r). \tag{28}$$

These relations provide important insight since they show that for infinitely diluted systems the pure solvent Eq. (26) is identical with the OZ Eq. (14), i.e. the presence of an infinitesimal trace amount of solute does not perturb the solvent, as expected. Consequentially, the uv Eq. (27) contains a pure solvent contribution that is *independent* of the kind of solute and that can therefore be precomputed for a given solvent model. This is evident by transforming the relation to reciprocal (k) space by a Fourier transform (as is common practice

for numerical solution to integral equation theories, but we will not further elaborate these methods here) which simplifies the convolution to a simple product:

$$\hat{h}^{uv}(k) = \hat{c}^{uv}(k) + \rho_0 \hat{c}^{uv}(k)\hat{h}^{vv}(k) = \hat{c}^{uv}(k)\,[1 + \rho_0 \hat{h}^{vv}(k)], \tag{29}$$

$$\rho_0 \hat{h}^{uv}(k) = \hat{c}^{uv}(k)\,[\rho_0 + \rho_0^2 \hat{h}^{vv}(k)] \equiv \hat{c}^{uv}(k)\,\hat{\chi}^{vv}(k). \tag{30}$$

Here, $\hat{\chi}^{vv}$ is the solvent susceptibility or density correlation function of the pure solvent since it describes physically the linear response function responsible for a density fluctuation in response to an external potential exerted on the solvent (in our case here due to the presence of a solute molecule).

Analogously, the uu Eq. (28) contains precomputable uv functions on the r.h.s., which allows for *non-iterative* and therefore very efficient calculations of solute-solute PMFs, for instance within the HNC approximation:

$$\begin{aligned} w^{uu}_{\text{HNC}}(r) &= -\beta^{-1} \ln g^{uu}_{\text{HNC}}(r) = u^{uu}(r) - \beta^{-1} t^{uu}_{\text{HNC}}(r) \\ &= u^{uu}(r) - \beta^{-1}(c^{uv} * \rho_0 h^{vu})\,(r). \end{aligned} \tag{31}$$

The second term in the last line reflects physically the solvation contribution to the total PMF.

In summary, a typical OZ study consists of several steps which are easily generalized to molecular situations described below: (a) solve the vv equation for a given solvent model, (b) use the resulting solvent susceptibility for different uv calculations representing different solute models, (c) solve the uu equations for various solute pairs, using their interaction potential and precomputed uv distribution functions. In this way, integral equation theory allows for the rapid access to the full set of distribution functions in complex solution systems, subject to the possibility to find convergent solutions to the nonlinear equations numerically and to apply adequate approximations to the bridge function (not discussed in detail here).

2.4 Chemical Potential and Free Energy from Integral Equations

An important next step after PDFs have been determined is the calculation of chemical potentials (and related Helmholtz free energies for the solvent by extending the coupling parameter scaling to all vv interactions) by inserting them into Eqs. (11)-(13). To this end, one has to account for the implicit coupling parameter dependence of all correlation functions appearing in conjunction with the scaled potential in the constitutive set of integral equation and closure, Eqs. (14) and (22). For the HNC approximation, this integration can indeed be performed directly by exploiting the fact that the HNC/OZ system yields an exact derivative for μ^{ex} and F which facilitates the derivation of a closed form that only depends on end point correlation functions[17,18], shown here for the chemical potential of a simple liquid:

$$\mu^{\text{ex}}_{\text{HNC}} = \rho_0 \int \mathbf{dr}\, \left[\frac{1}{2}h^2(r) - c(r) - \frac{1}{2}h(r)\,c(r)\right]. \tag{32}$$

Note that TCFs and DCFs appearing in this relation can be vv or uv functions; in the former case one would therefore obtain the excess chemical potential of the pure solvent.

Problematic issues arise if one assumes non-zero bridge functions, for which the coupling parameter dependence of B has to be known. For instance, if B is given as a function of the ICF t, as within the PY approximation, Eq. (24), the λ dependence of c and h would implicitly scale B. However, it turned out that such a bridge definition violates an essential principle of thermodynamics, the independence of a state function such as the free energy from the integration path[19]. This problem arises also in cases where the bridge function is directly extracted from simulation data where the scaling behavior is unknown, and for practically all classes of bridge models and free energy functionals in use (for an overview of available free energy models see[20] and references cited therein). One way to cope with the situation is to invoke a perturbation expansion where an end point (i.e. fully scaled) bridge function enters into a correction to the HNC chemical potential[21].

The physically more rigorous approach involves searching for properties of the bridge function that violate or control the path independence. To this end, a formalism based on a general variational principle that is equivalent to the path independence demanded by thermodynamics has been developed by us[22]. Without going into details, it could be shown that several functional classes of bridge function dependencies on correlation functions can indeed be constructed in a path-independent manner, most notably if the bridge dependence is given by a "renormalized" ICF, .i.e. by writing $B(r) = B(t(r) - \beta u(r)) \equiv B(t^*)$. In this way, it is possible to derive closed form expressions for the chemical potential for closure approximations that satisfy certain numerical stability issues associated with HNC, termed "partial series expansion" closure of order n (PSE-n)[23]:

$$\mu^{\text{ex}}_{\text{PSE-}n} = \rho_0 \int \mathrm{d}\mathbf{r} \left[\frac{1}{2} h^2(r) - c(r) - \frac{1}{2} h(r)\, c(r) - \Theta(h(r)) \frac{(t^*(r))^{k+1}}{(k+1)!} \right], \tag{33}$$

again for simple liquids. All derivations made so far directly apply also to the molecular case discussed in the next chapter, requiring only straightforward generalizations by summing over site components as defined in Eqs. (12) and (13).

An important physical aspect is associated with path-independent free energy or chemical potential functionals, namely its stationarity with respect to variation of correlation functions[18,23]. This leads essentially to the compact expression

$$\frac{\delta \mu^{\text{ex}}}{\delta u(\mathbf{r})} = \rho_0 g(\mathbf{r}) \tag{34}$$

for simple liquids (again with straightforward molecular generalization). In practice, such a property is useful for two reasons. Firstly, it allows for the calculation of gradients of free energy surfaces with respect to nuclear coordinates $\mathbf{R}$, in a generic formulation given by

$$\nabla_{\mathbf{R}} \mu^{\text{ex}}(\mathbf{R}) = \rho_0 \int \mathrm{d}\mathbf{r}\, g(\mathbf{r};\mathbf{R})\, \nabla_{\mathbf{R}} U(\mathbf{r};\mathbf{R}). \tag{35}$$

This result allows for geometry optimization in solution and coupling of an integral equation framework with MD simulations in the sense of an implicit description of solvent degrees of freedom[24]. Secondly, stationarity of the chemical potential can be coupled with the quantum-mechanical variational principle to construct a total stationary functional for computing self-consistent solute wave functions in solution (see below).

3 Molecular Theories

3.1 Reference Interaction Site Model

In comparison with simple liquids composed of spherically symmetric particles, molecular solution systems pose additional challenges due to the dimensionality problem. For instance, the description of the correlation between two anisotropic rigid bodies requires six coordinates, three for relative distance and three for angular degrees of freedom. Although the MOZ equation[11] is a conceptually straightforward generalization of integral equation theories of simple liquids, its numerical solution on a six-dimensional (6D) grid is technically extraordinarily demanding and has not been attempted yet. Instead, an expansion of the correlation functions in a set of suitable basis functions (rotational invariants) is possible. Coupled with efficient numerical schemes, this approach is nevertheless mathematically and algorithmically challenging and has found only limited, but fascinating application, e.g.[25].

It is therefore advisable to search for simpler schemes that yield site-site distribution functions (scaling with the product of the number of solute and solvent sites, and the number of radial grid points) or solute molecule / solvent site PDFs (scaling with the number of solvent sites and of 3D grid points). Starting with the MOZ equation, this involves an approximation of the 6D molecule-molecule DCF in terms of simpler site-site or molecule-site functions. The first approach leads to the 1D "reference interaction site model" (RISM) integral equation theory (also termed site-site OZ, SSOZ)[26,27], the second to the 3D RISM methodology[28,29,30]. Conceptually, the 6D DCF is approximated by a sum of site-site or molecule-site components, leading to the possibility to integrate out the angular dependence which leads to the occurrence of *intramolecular* site-site correlation functions for specifying the solvent alone (3D RISM) or the solute and the solvent (1D RISM).

In summary, the 1D RISM equations are given in matrix formulation by (see also refs.[2,14,15])

$$\rho_0 \mathbf{h}^{vv} = \omega^v * \mathbf{c}^{vv} * (\rho_0 \omega^v + \rho_0^2 \mathbf{h}^{vv}), \tag{36}$$

$$\rho_0 \mathbf{h}^{uv} = \omega^u * \mathbf{c}^{uv} * (\rho_0 \omega^v + \rho_0^2 \mathbf{h}^{vv}) = \omega^u * \mathbf{c}^{uv} * \chi^{vv}, \tag{37}$$

$$\mathbf{h}^{uu} = \omega^{u(1)} * \mathbf{c}^{uu} * \omega^{u(2)} + \omega^{u(1)} * \mathbf{c}^{u(1)v} * \rho_0 \mathbf{h}^{u(2)v} \tag{38}$$

for single-component solvents, where each matrix element represents a solute site / solvent-site pair, e.g. $h_{\alpha\gamma}(r)$. The intramolecular correlation matrix ω consists of elements that, for rigid molecules, are essentially Dirac δ-functions that constrain site distances to fixed $l_{\alpha\alpha'}$ or $l_{\gamma\gamma'}$, represented in reciprocal space by

$$\hat{\omega}(k) = \frac{\sin(kl)}{kl}. \tag{39}$$

One immediately sees that the hierarchy of correlation functions defined in section 2.3 still holds by considering each site as a component of a mixture. For dielectric solvents, the equations have to be modified to a dielectrically consistent (DRISM) variant[31,32], since the 1D RISM dielectric permittivity yields only a trivial result[7]. It is important to note that the 1D RISM approximation, if used in conjunction with a generic closure similar to Eq. (22), contains well characterized artifacts originating from the fact that Eq. (23) does not hold,

leaving the bridge function only as a formal quantity. This also becomes apparent from a derivation within a cDFT framework[33].

The 3D RISM equations generically result from setting the intramolecular correlation matrix of the solute to a scalar with value one in reciprocal space, leaving the vv part as before for 1D RISM. We then obtain, accounting for an arbitrary shape of the solute by changing the radial distance to a 3D grid coordinate,

$$\rho_0 \mathbf{h}^{uv} = \mathbf{c}^{uv} * \chi^{vv}, \tag{40}$$

$$\mathbf{h}^{uu} = \mathbf{c}^{uu} + \mathbf{c}^{u(1)v} * \rho_0 \mathbf{h}^{u(2)v}, \tag{41}$$

where in the last equation the solute pair is placed in a fixed relative orientation[16]. While still hampered by approximation errors due to the site-site solvent description, the 3D RISM approach represents the most successful, technically manageable, and widely adopted integral equation methodology today. One has to note though that, related to approximation shortcomings, relative free energies are more reliable than absolute ones due to error compensation.

3.2 Coupling RISM Theory with a Polarizable Solute Hamiltonian

Thus far, we assumed a given solute-solvent interaction potential that does not change upon immersing a solute molecule in a possibly polar solvent, reflecting an unpolarizable model. Physical reality, however, dictates changes of the solute's charge distribution when it is exposed to a polarizing environment. This is not only true for dipolar solvents which certainly exhibit the strongest effects, but also for fluids with non-zero higher electric multipole moments. If the solute is described by a classical fixed charge force field, polarization effects can only be modeled to some extent if different conformations exhibit varying charge distributions, but will definitely fail if polarization is essential for quantitative predictions of solvation free energies. To this end, solvent-induced electric fields have to be coupled to an adequately responding charge density; of course in principle bidirectionally by repolarizing single solvent molecules, but this latter process is very hard to describe statistical-mechanically.

Both, a quantum-mechanical description of the solute's charge density and a classically polarizable force field can be used to serve the purpose to add more physical reality. As to the former, the polarizable continuum model[34] (PCM) is the most widely adopted methodology to couple a continuum solvent model in the sense of Poisson-Boltzmann theory to an electronic wave function. In this, as in other implicit solvation approaches, one has to keep in mind the limitations of the thermodynamic framework. The wave function ψ experiences an equilibrium solvent, which implies that we assume for an observable associated with operator $\hat{O}$

$$\left\langle \langle\psi|\hat{O}|\psi\rangle \right\rangle_T \approx \left\langle \langle\psi\rangle_T|\hat{O}|\langle\psi\rangle_T \right\rangle, \tag{42}$$

i.e. a (true) observable that is thermally averaged (subscript "T") over solvent configurations is taken to be identical to one computed with a thermally averaged wave function (defined as being exposed to an equilibrium solvent). This is of course only an approximation that may fail dramatically in certain situations. Even worse are phenomena where

nonequilibrium configurations play a role, most notably for calculations of spectroscopic properties.

If we, however, accept the approximation, we need a criterion from which a joint, self-consistent electronic and solvent structure can be determined. We can again refer to Eq. (5) to this end by defining an "absolute" free energy of a solute in solution via

$$F = \mu^{\text{ex}} + E_{\text{sol}} \tag{43}$$

where E_{sol} represents the quantum-mechanical solute energy in solution (taken as an intensive, "per particle" quantity as is also assumed for F). Note that F can be safely replaced by the Gibbs energy G if the solvation process at constant pressure is hypothetically performed at a constant solvent volume whose average corresponds to this pressure[35]. This ansatz allows for the construction of the target criterion, namely a variational principle that minimizes F upon simultaneous variation of wave function and chemical potential. While the former is a fundamental concept in quantum mechanics, the latter is valid for closures that satisfy stationarity with respect to variation of solvent structure, as defined by correlation functions, see Eq. (34). This forms the basis of the RISM-SCF (self-consistent field) approach developed for 1D and 3D setups as well as for its variants including an MOZ implementation (for an overview, see[36] and references cited therein).

An alternative that does not require global variational principles is provided by an approach adopted from "cluster embedding" strategies developed originally for solid state environments. Here, one attempts to model the effect of a periodic crystal on a small, elementary fragment by a set of embedding point charges that are fitted to represent the true electrostatic field at the location of a cluster of atoms. Such a strategy can be immediately adapted to a 3D RISM solvation framework, a methodology developed and termed ("embedding cluster") EC-RISM by us[37]. Briefly, we identify embedding point charges q at grid point $\mathbf{r}_i$ with the locally integrated charge density ρ_q (that originates from a solvent site charge q_γ) over the grid point volume V_i,

$$q(\mathbf{r}_i) = \int_{V_i} \mathrm{d}\mathbf{r} \sum_\gamma q_\gamma \rho_{0,\gamma} g_\gamma(\mathbf{r}) \approx \rho_q(\mathbf{r})\, V_i. \tag{44}$$

The embedding charges polarize the electronic solute Hamiltonian, giving rise to an update of the solute-solvent interaction potential (calculated approximately from solute point charges or, exactly, from the electrostatic potential determined by the wave function), that leads in turn to a new set of correlation functions and resulting grid charges. The cycle is repeated until mutual self-consistency between solute electronic and solvent structure is achieved.

In practice, the EC-RISM methodology works well for a variety of cases, most recently for the prediction of nuclear magnetic resonance (NMR) magnetic shieldings and chemical shifts (see[38] and references therein for other applications). The essentially same strategy can also be employed if the charge relaxation is not modeled by an electronic wave function, but by a classically polarizable solute force field[39]. In this case, a change of atomic polarizabilities in response to a polarizing environment is also initiated by a set of embedding point charges, yielding an updated electrostatic potential on the grid which is in turn inserted into the closure relation for computing a new set of correlation functions. As for quantum-mechanical EC-RISM applications, a typically unpolarizable, fixed charge solvent force field is combined with a polarizable solute model.

3.3 Selected Applications

As an illustrative example taken from recent work performed in the group of the authors we show practical calculation results for aqueous solutions in order to demonstrate advantages and limitations of typical integral equation applications.

Water is of course the most relevant solvent for all kinds of chemical and biochemical processes. Therefore, much work has been invested into developing water models suitable for MD simulations, i.e. with as few details as necessary to describe structure, dynamics, and thermodynamics well. These, for instance, three-site models such as TIP3P or SPC/E[40] can also be used in 1D RISM applications with one notable extension, namely that non-zero Lennard-Jones (LJ) parameters have to be assigned to hydrogen in order to prevent divergence of the numerical solution.

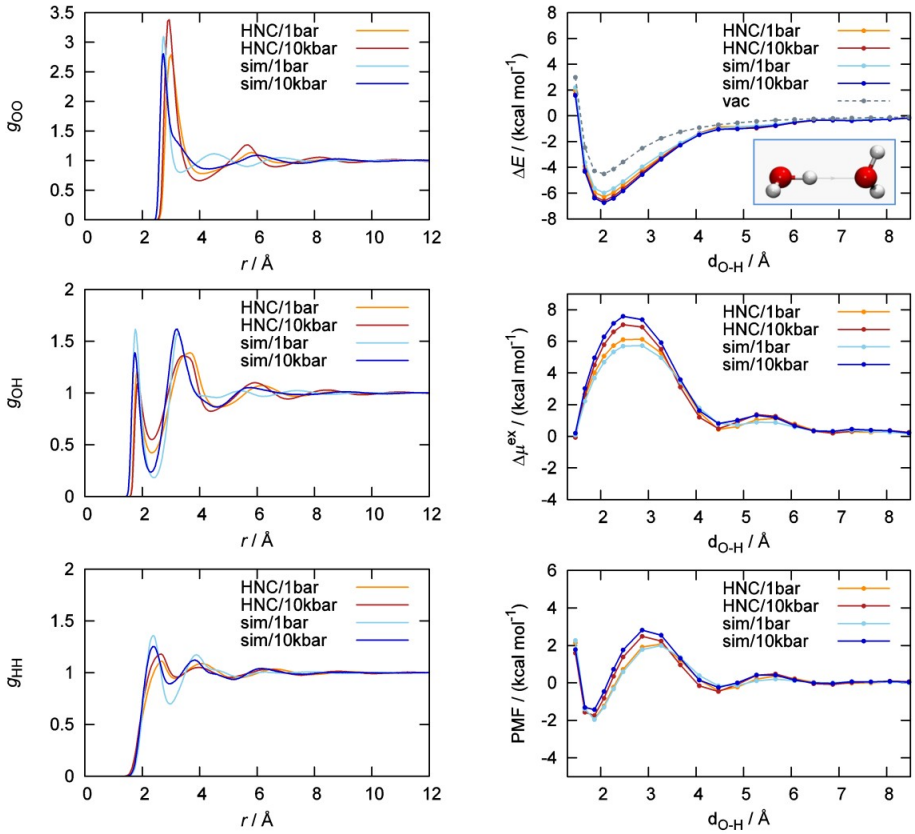


Figure 1: Left: site-site PDFs for liquid SPC/E water under various pressures from 1D RISM/HNC and from MD simulations; right: water dimer PMF in water and its components, calculated with solvent susceptibilites taken from 1D RISM/HNC or from MD (reference simulations by D. Horinek, Regensburg).

An interesting research area is the study of solution systems under extreme conditions, particularly high pressure, since proteins tend to explore entirely different high-energy,

typically unfolded conformational regions under these conditions. Understanding extreme environmental condition properties also serves to possibly learn about the origin of life or fundamentals of exobiology. Therefore, a key question is to examine the impact of high pressure on water structure and PMFs of model systems.

Figure 1 shows results of 1D and 3D RISM calculations on SPC/E water and a water dimer immersed in water (for 1D RISM, hydrogen LJ parameters were assigned from earlier work[39]). One can see that RISM results with the HNC closure are semiquantitatively correct with notable exceptions concerning oxygen-oxygen packing structure (which in turn can be improved by smaller hydrogen LJ parameters at the prize of worsening H-H and O-H structure), particularly at high pressure. It is, however possible to directly extract the solvent site-site susceptibility χ from MD simulations[41] which basically corrects for 1D RISM deficits and eliminates a source of error for subsequent 3D RISM calculations. A comparison of both approaches (χ taken from 1D RISM/HNC and from MD within EC-RISM) is also shown in Figure 1 for the water dimer. Figure 2 shows corresponding 3D oxygen densities for the dimer at different distances for χ from MD, illustrating the effect of high pressure on the solvation shell structure which becomes more pronounced. We also identify the origin of the PMF barrier at close distance, the displacement of water molecules between the pair.

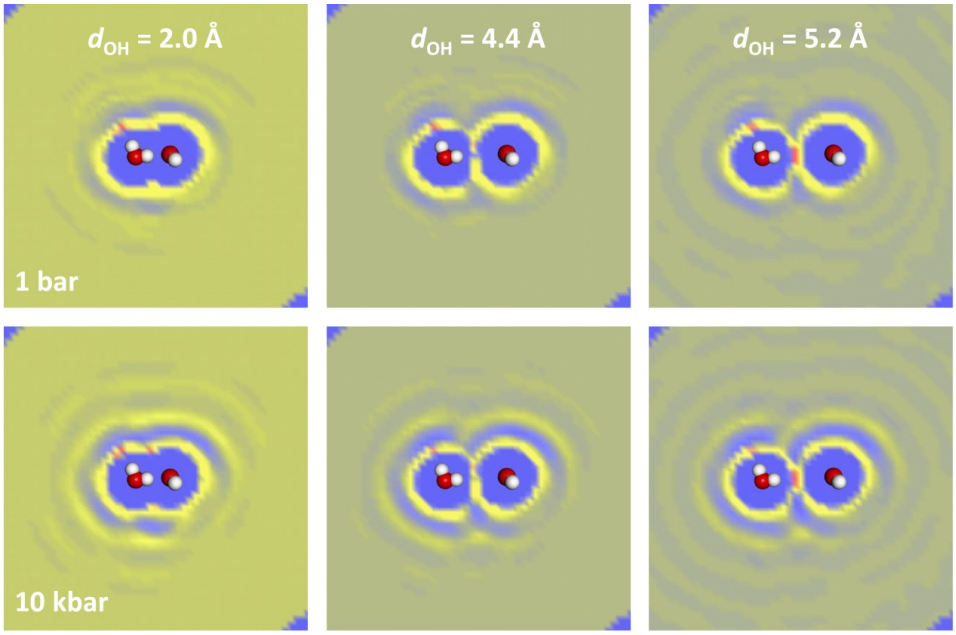


Figure 2: 3D oxygen distribution around a water dimer for various distances at ambient and high pressure, computed using EC-RISM/PSE-2/MP2/6-311++G(2d,2p). Blue corresponds to solvent-depleted regions, yellow to red to enrichments.

In Figure 1, we can furthermore observe a typical feature of RISM calculations, error compensation in thermodynamic quantities. This is due to the fact that the PMF contains

the isolated water molecules as reference state, treated with the same level of theory. As a consequence, the 1D RISM level of theory does not play a substantial role for features of the PMF. Both approaches yield closely corresponding curves that agree even better at ambient pressure.

An interesting feature of a charge distribution that is allowed to relax due to environmental polarization can also be identified from Figure 1: Even though the separate components, electronic energy and chemical potential in solution are quite sensitive towards pressure perturbation, the total PMF apparently is not. This is almost perfectly satisfied at 1 bar pressure, i.e. at conditions used for typical force field applications. In principle, we see here the balance between strain on the wave function and more favorable chemical potential.

4 Concluding Remarks

While we have shown that integral equation theories as a liquid state theory with sound physical and largely non-empirical motivation can be successfully employed in a variety of situations, many question have to be left unanswered. From our point of view, the most urgent issue concerns the construction of bridge functions that provide good quality approximations to liquid structure and thermodynamics at the same time. This has not been achieved with present day approaches and represents a major conceptual and numerical challenge. Work in our laboratory is ongoing to address these questions in the context of developing software for numerical solutions. Both questions, physics and numerical mathematics, have to be viewed in context, since a physically correct theory should also have numerically stable solutions, if they exist.

Acknowledgements

We have to thank a number of developers, contributors, and users of the software: T. Kloss, S. Tayefeh, N. Tschauder, K. Kuhn, M. Kreim, S. Ehrhart, R. Frach, D. Tomazic, L. M. Henkes, F. Mrugalla, P. Kibies, M. Urban, S. Egbers, C. Schulz, S. Strohfeldt, K. F. Schmidt, S. Güssregen. Financial support by the Bundesministerium für Bildung und Forschung (BMBF) within the *hpCADD* project is acknowledged as well as by the Deutsche Forschungsgemeinschaft (DFG), Stiftung Mercator, and Sanofi.

References

1. M. K. Gilson, J. A. Given, B. L. Bush, and J. A. McCammon, *The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review*, Biophys. J. **72**, 1047-1069, 1997.
2. K. F. Schmidt and S. M. Kast, *Hybrid Integral Equation/Monte Carlo Approach to Complexation Thermodynamics*, J. Phys. Chem. B **106**, 6289-6297, 2002.
3. B. Roux and T. Simonson, *Implicit Solvent Models*, Biophys. Chem. **78**, 1-20, 1999.
4. D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).

5. J. G. Kirkwood, *Statistical Mechanics of Fluid Mixtures*, J. Chem. Phys. **3**, 300-313, 1935.
6. P. Kollman, *Free Energy Calculations: Applications to Chemical and Biochemical Phenomena*, Chem. Rev. **93**, 2395-2417, 1993.
7. J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 3rd ed. (Academic Press, Oxford, 2006).
8. R. Evans, *The Nature of the Liquid-Vapour Interface and Other Topics in the Statistical Mechanics of Non-uniform, Classical Fluids*, Adv. Phys. **28**, 143-200, 1979.
9. R. Ramirez, R. Gebauer, M. Mareschal, and D. Borgis, *Density Functional Theory of Solvation in a Polar Solvent: Extracting the Functional from Homogeneous Solvent Simulations*, Phys. Rev. E **66**, 031206, 2002.
10. J. D. McCoy and A. D. J. Haymet, *Theory of Freezing: The Inhomogeneous Ornstein-Zernike Equation*, Int. J. Thermophys. **10**, 87-100, 1989.
11. L. Blum and A. J. Torruella, *Invariant Expansion for Two-Body Correlations: Thermodynamic Functions, Scattering, and the Ornstein-Zernike Equation*, J. Chem. Phys. **56**, 303-310, 1972.
12. M. Oettel, *Integral Equations for Simple Fluids in a General Reference Functional Approach*, J. Phys.: Condens. Matter **17**, 429-452, 2005.
13. T. Morita and K. Hiroike, *Integral Equation for Pair Distribution Function*, Theor. Phys. **23**, 385-387, 1960.
14. F. Hirata, P. J. Rossky, and B. M. Pettitt, *The Interionic Potential of Mean Force in a Molecular Polar Solvent from an Extended RISM Equation*, J. Chem. Phys. **78**, 4133-4144, 1983.
15. B. M. Pettitt and M. Karplus, *The Potential of Mean Force between Polyatomic Molecules in Polar Molecular Solvents*, J. Chem. Phys. **83**, 781-789, 1985.
16. A. Kovalenko and F. Hirata, *Potential of Mean Force between Two Molecular Ions in a Polar Molecular Solvent: A Study by the Three-Dimensional Reference Interaction Site Model*, J. Phys. Chem. B **103**, 7942-7957, 1999.
17. T. Morita and K. Hiroike, *A New Approach to the Theory of Classical Fluids. I*, Prog. Theor. Phys. **23**, 1003-1027, 1960.
18. S. J. Singer and D. Chandler, *Free Energy Functions in the Extended RISM Approximation*, Molec. Phys. **55**, 621-625, 1985.
19. R. Kjellander and S. Sarman, *The Chemical Potential of Simple Fluids in a Common Class of Integral Equation Closures*, J. Chem. Phys. **90**, 2768-2775, 1989.
20. H. Freedman, L. Le, J. A. Tuszynski, and T. N. Truong, *Improving the Performance of the Coupled Reference Interaction Site Model-Hyper-netted Chain (RISM-HNC) / Simulation Method for Free Energy of Solvation*, J. Phys. Chem. B **112**, 2340-2348, 2008.
21. A. Kovalenko and F. Hirata, *Hydration Free Energy of Hydrophobic Solutes Studied by a Reference Interaction Site Model with a Repulsive Bridge Correction and a Thermodynamic Perturbation Method*, J. Chem. Phys. **113**, 2793-2805, 2000.
22. S. M. Kast, *Free Energies from Integral Equation Theories: Enforcing Path Independence*, Phys. Rev. E **67**, 041203, 2003.
23. S. M. Kast and T. Kloss, *Closed-Form Expressions of the Chemical Potential for Integral Equation Closures with Certain Bridge Functions*, J. Chem. Phys. **129**, 236101, 2008.

24. T. Luchko, S. Gusarov, D. R. Roe, C. Simmerling, D. A. Case, J. Tuszynski, and A. Kovalenko, *Three-Dimensional Molecular Theory of Solvation Coupled with Molecular Dynamics in Amber*, J. Chem. Theor. Comput. **6**, 607-624, 2010.
25. J. Richardi, C. Millot, and P. H. Fries, *A Molecular Ornstein-Zernike Study of Popular Models for Water and Methanol*, J. Chem. Phys. **110**, 1138-1147, 1999.
26. D. Chandler and H. C. Andersen, *Optimized Cluster Expansions for Classical Fluids. I. General Theory and Variational Formulation of the Mean Spherical Model and Hard Sphere Percus-Yevick Equations*. J. Chem. Phys. **57**, 1918-1929, 1972.
27. D. Chandler and H. C. Andersen, *Optimized Cluster Expansions for Classical Fluids. II. Theory of Molecular Liquids*. J. Chem. Phys. **57**, 1930-1937, 1972.
28. M. Ikeguchi and J. Doi, *Direct Numerical Solution of the Ornstein–Zernike Integral Equation and Spatial Distribution of Water around Hydrophobic Molecules*, J. Chem. Phys. **103**, 5011-5017, 1995.
29. D. Beglov and B. Roux, *An Integral Equation to Describe Solvation of Polar Molecules in Liquid Water*, J. Phys. Chem. B **101**, 7821-7826, 1997.
30. A. Kovalenko and F. Hirata, *Three-dimensional Density Profiles of Water in Contact with a Solute of Arbitrary Shape: A RISM Approach*, Chem. Phys. Lett. **290**, 237-244, 1998.
31. J. S. Perkyns and B. M. Pettitt, *A Dielectrically Consistent Interaction Site Theory for Solvent—Electrolyte Mixtures*, Chem. Phys. Lett. **190**, 626-630, 1992.
32. J. S. Perkyns and B. M. Pettitt, *A Site-Site Theory for Finite Concentration Saline Solutions*, J. Chem. Phys. **97**, 7656-7666, 1992.
33. D. Chandler, J. D. McCoy, and S. J. Singer, *Density Functional Theory of Nonuniform Polyatomic Systems. II. Rational Closures for Integral Equations*, J. Chem. Phys. **85**, 5977-5982, 1986.
34. J. Tomasi, B. Mennucci, and R. Cammi, *Quantum Mechanical Continuum Solvation Models*, Chem. Rev. **105**, 2999-3094, 2005.
35. S.-H. Chong and S. Ham, *Thermodynamic-Ensemble Independence of Solvation Free Energy*, J. Chem. Theor. Comput. **11**, 378-380, 2015.
36. H. Sato, *A Modern Solvation Theory: Quantum Chemistry and Statistical Chemistry*, Phys. Chem. Chem. Phys. **15**, 7450-7465, 2013.
37. T. Kloss, J. Heil, and S. M. Kast, *Quantum Chemistry in Solution by Combining 3D Integral Equation Theory with a Cluster Embedding Approach*, J. Phys. Chem. B **112**, 4337-4343, 2008.
38. R. Frach and S. M. Kast, *Solvation Effects on Chemical Shifts by Embedded Cluster Integral Equation Theory*, J. Phys. Chem. A **118**, 11620-11628, 2014.
39. F. Hoffgaard, J. Heil, and S. M. Kast, *Three-Dimensional RISM Integral Equation Theory for Polarizable Solute Models*, J. Chem. Theory Comput. **9**, 4718-4726, 2013.
40. B. Guillot, *A Reappraisal of What we have Learnt During Three Decades of Computer Simulation of Water*, J. Mol. Liquids **101**, 219-260, 2002.
41. S. M. Kast, K. F. Schmidt, and B. Schilling, *Integral Equation Theory for Correcting Truncation Errors in Molecular Simulations*, Chem. Phys. Lett. **367**, 398-404, 2003.

Smoothed Dissipative Particle Dynamics - A Mesoscopic Particle-Based Hydrodynamic Technique for Complex Fluids

Dmitry A. Fedosov, Kathrin Müller, and Gerhard Gompper

Theoretical Soft Matter and Biophysics, Institute of Complex Systems
and
Institute for Advanced Simulation
Forschungszentrum Jülich, 52425 Jülich, Germany

E-mail: d.fedosov@fz-juelich.de

Smoothed dissipative particle dynamics (SDPD) combines two popular mesoscopic techniques, the smoothed particle hydrodynamics and dissipative particle dynamics (DPD) methods, and can be considered as an improved DPD approach. Advantages of the SDPD method over conventional DPD include the possibility of using an arbitrary equation of state, direct input of transport properties, and a well-defined physical scale of discretized elements or fluid particles. The SDPD method has been already applied to a number of mesoscopic problems involving complex fluids. Despite several advantages of the SDPD method over the conventional DPD model, the original formulation of SDPD by Español and Revenga (2003) lacks angular momentum conservation, leading to unphysical results for problems where the conservation of angular momentum is essential. To overcome this limitation, the SDPD method is extended by introducing a particle spin variable such that local and global angular momentum conservation is restored. The new SDPD formulation (SDPD+a) is directly derived from the Navier-Stokes equation for fluids with spin, while thermal fluctuations are incorporated similar to the DPD method. In this chapter, we describe the basics of the SDPD method and show several of its applications including particle margination in blood flow. In addition, the SDPD method with angular momentum conservation is validated using two problems: (i) the Taylor-Couette flow with two immiscible fluids and (ii) a tank-treading vesicle in shear flow with a viscosity contrast between inner and outer fluids. For both problems, the new SDPD method leads to simulation predictions in agreement with the corresponding analytical theories, while the original SDPD method fails to capture properly physical characteristics of the systems due to violation of angular momentum conservation.

1 Introduction

Mesoscopic hydrodynamic simulations, such as the Lattice-Boltzmann (LB) method[1], dissipative particle dynamics (DPD)[2–4], multi-particle collision dynamics (MPC)[5,6], smoothed particle hydrodynamics (SPH)[7,8] etc., are frequently used to investigate a wide range of problems including colloidal and polymer solutions, dynamics of microswimmers, tissue growth, and flow behavior of vesicles and cells. All these examples include mesoscopic length scales (e.g., the size of suspended particles) rendering the modeling on atomistic level impossible. A continuum approximation is also not appropriate for such problems due to the loss of necessary mesoscopic details. Thus, large scientific efforts have been invested to derive reliable and efficient mesoscopic simulation techniques, which are able to tackle a wide range of problems.

A recently established mesoscopic method, smoothed dissipative particle dynamics (SDPD)[9], combines advantages of two popular techniques namely SPH and DPD. The

SDPD method for fluid flow is directly derived from the Navier-Stokes equation using a Lagrangian discretization similar to SPH, while the inclusion of thermal fluctuations in SDPD is similar to that in the DPD formalism. SDPD can also be considered as an improved DPD method. Advantages of the SDPD method over conventional DPD include the possibility of using an arbitrary equation of state, direct input of transport properties, and a well-defined physical scale of discretized elements or fluid particles. In addition, it has been shown that the SDPD method produces proper scaling of thermal fluctuations for different fluid particle sizes[10]. The SDPD method has been already applied to a number of problems including simulations of different particles[11] and polymers[12] in a suspension, single red blood cells in tube flow[13], margination of leukocytes[14], and margination of different particles[15] in blood flow.

Despite the advantages of SDPD over the DPD method, the original SDPD formulation[9] does not conserve angular momentum, both locally and globally. Recent numerical simulations using the MPC method[16] have shown that angular momentum conservation is essential in some problems including Taylor-Couette flow with two immiscible fluids and vesicle tank-treading in shear flow. A violation of angular momentum conservation may lead to an asymmetric stress tensor and spurious unphysical torques, resulting in erroneous simulation results. To derive a consistent version of SDPD with angular momentum conservation, we introduce a spin variable, such that each SDPD particle possesses an angular velocity[17]. This idea is similar to that of the fluid particle (FPM) model[18], where every particle possesses an angular velocity; however, FPM lacks a direct connection to the discretization of the Navier-Stokes equation. Also, a spin variable has been introduced in the single-particle DPD formulation[19], where a colloidal particle can be represented by a single DPD particle with spin. Consistent SDPD formulation with angular momentum conservation is obtained by a direct discretization of the Navier-Stokes equation for a fluid with spin[20].

We will show several applications of the SDPD method. The SDPD method will be applied to the problem of micro- and nano-particle margination (i.e., particle migration toward the walls) in blood flow, which is a crucial step in drug delivery since it is a precondition for particle adhesion at the vessel walls. The importance of angular momentum conservation will be illustrated using two examples. First, the Taylor-Couette flow with two immiscible fluids is simulated showing that the extended SDPD method results in predictions in agreement with the analytical solution derived from the Navier-Stokes equation. The second problem to test the SDPD method with angular momentum conservation is a tank-treading vesicle in shear flow, which has been described theoretically by Keller and Skalak[21]. Vesicle tank-treading in shear flow corresponds to rotational motion of a membrane around the vesicle center-of-mass, while the vesicle preserves its stationary shape with a finite inclination angle. The new SDPD formulation results in predictions of vesicle inclination angles for several viscosity contrasts between inner and outer fluids in agreement with the Keller-Skalak theory[21], while the SDPD method without angular momentum conservation clearly fails to capture quantitatively correct dynamics.

2 Smoothed Particle Hydrodynamics

The basic equations to describe viscous fluid flow are the Navier-Stokes (NS) equations[20]. In a Lagrangian description, the NS equations take the form

$$\rho \frac{d\mathbf{v}}{dt} = -\nabla p + \eta \nabla^2 \mathbf{v} + \left(\frac{\eta}{3} + \xi\right) \nabla \nabla \cdot \mathbf{v}, \tag{1}$$

where $\mathbf{v}$ is the fluid's velocity, ρ is the fluid's density, p is the pressure, η is the dynamic shear viscosity, and ξ is the bulk viscosity. For the closure of the equations above, we also need to add the continuity equation given by

$$\frac{d\rho}{dt} + \rho \nabla \cdot \mathbf{v} = 0. \tag{2}$$

The smoothed particle hydrodynamics (SPH) method[7,22] for fluid flow is a Lagrangian discretization of Eqs. (1) and (2). In the SPH method, a field variable $\tilde{g}(\mathbf{r})$ is replaced by the convolution integral of a field $g(\mathbf{r})$ and a kernel function $W(\mathbf{r}, h)$ as,

$$\tilde{g}(\mathbf{r}) \approx \int_V g(\mathbf{r}') W(\mathbf{r} - \mathbf{r}', h)\, dV', \tag{3}$$

where the kernel function has to be differentiable and depends on the distance $|\mathbf{r} - \mathbf{r}'|$ and the smoothing length h. In addition, the integral over $W(\mathbf{r} - \mathbf{r}', h)$ has to be normalized and the condition $\lim_{h \to 0} W(\mathbf{r} - \mathbf{r}', h) = \delta(\mathbf{r} - \mathbf{r}')$ needs to be satisfied. For $W(\mathbf{r}, h)$ being the delta function, we would have $\tilde{g}(\mathbf{r}) = g(\mathbf{r})$. The convolution integral is discretized using small fluid volumes (or particles) such that $\rho(\mathbf{r}')dV' \to m_j$ with m_j being the mass and $\rho(\mathbf{r}') \to \rho(\mathbf{r}_j)$ being the mass density of particle j at the position vector $\mathbf{r}_j$. The discretized convolution integral is then given by

$$\tilde{g}(\mathbf{r}_i) \approx \sum_{j=1}^{N} \frac{m_j}{\rho(\mathbf{r}_j)} g(\mathbf{r}_j) W(|\mathbf{r}_i - \mathbf{r}_j|, h), \tag{4}$$

where N is the number of particles (Lagrangian discretization points) within the volume V characterized by the smoothing radius h. Furthermore, derivatives of the field variable $g(\mathbf{r})$ follow similar approximation strategy which is described in Appendix A. Further in the text, we will also use the notations $\rho(\mathbf{r}_j) = \rho_j$, $g(\mathbf{r}_j) = g_j$, and $W(|\mathbf{r}_i - \mathbf{r}_j|, h) = W_{ij}$.

Using the SPH formalism, the continuity equation (2) becomes (see Eq. (31))

$$\frac{d\rho_i}{dt} = \sum_j m_j \mathbf{v}_{ij} \cdot \nabla_i W_{ij}, \tag{5}$$

where $\nabla_i W_{ij}$ can be analytically calculated. The particle density ρ_i is defined as

$$\rho_i = \sum_j m_j W_{ij}. \tag{6}$$

Hence, the density of particle i can be computed using its neighboring particles located within a sphere with a radius h. Similarly, different terms of the NS equation (1) can be discretized to obtain the equations which govern particle dynamics. Using the Newton's

second law of motion $m_i d\mathbf{v}_i/dt = \mathbf{F}_i$ and the rules in Eqs. (32)-(36) of Appendix A we obtain the two forces: conservative (C) and dissipative (D) given by

$$\mathbf{F}_{ij}^C = \left(\frac{p_i}{\rho_i^2} + \frac{p_j}{\rho_j^2}\right) F_{ij}\mathbf{r}_{ij},$$
$$\mathbf{F}_{ij}^D = -\left(\frac{5\eta}{3} - \xi\right)\frac{F_{ij}}{\rho_i\rho_j}\mathbf{v}_{ij} - 5\left(\frac{\eta}{3} + \xi\right)\frac{F_{ij}}{\rho_i\rho_j}\hat{\mathbf{e}}_{ij}(\hat{\mathbf{e}}_{ij}\cdot\mathbf{v}_{ij}), \tag{7}$$

where p_i is the particle pressure, $\hat{\mathbf{e}}_{ij} = \mathbf{r}_{ij}/|\mathbf{r}_{ij}|$. A function $F(\mathbf{r}_{ij}) = F_{ij} \geq 0$ is defined such that $\nabla_i W_{ij} = -\mathbf{r}_{ij}F_{ij}$, and $\mathbf{F}_i = \sum_j \left(\mathbf{F}_{ij}^C + \mathbf{F}_{ij}^D\right)$, where the sum runs over all neighboring particles j of the particle i within the radius h. The conservative force controls locally the pressure field in the system. The dissipative force provides translational friction leading to the reduction of the velocity difference between two particles.

Time evolution of the position and the velocity of a particle i follows the Newton's second law as

$$\dot{\mathbf{r}}_i = \mathbf{v}_i, \quad \dot{\mathbf{v}}_i = \mathbf{F}_i = \sum_j \frac{1}{m_j}\mathbf{F}_{ij}. \tag{8}$$

Equation (8) is integrated using the velocity-Verlet algorithm[23]. Finally, in simulations the Lucy function[7]

$$W(r) = \frac{105}{16\pi h^3}\left(1 + 3\frac{r}{h}\right)\left(1 - \frac{r}{h}\right)^3, \tag{9}$$

is often used as a kernel function, which leads to $F(r) = \frac{315}{4\pi h^5}\left(1 - \frac{r}{h}\right)^2$. The equation of state for the pressure is often chosen to be

$$p = p_0\left(\frac{\rho}{\rho_0}\right)^\alpha + b, \tag{10}$$

where ρ_0 is the reference density, and the parameters p_0, α, and b can be freely selected. This pressure equation yields the speed of sound $c^2 = p_0\alpha/\rho_0$, which can be easily controlled through the above parameters resulting in a good approximation of fluid incompressibility[22,24].

3 Smoothed Dissipative Particle Dynamics

The smoothed dissipative particle dynamics (SDPD) method proposed by Español and Revenga[9] is a mesoscopic particle-based hydrodynamic approach which has been derived from the SPH[7,22] and DPD[2,3] simulation methods. Thus, in addition to the hydrodynamics described by SPH, consistent thermal fluctuations (i.e., they satisfy a balance between dissipative and random contributions) similar to those in the DPD method have to be added. To derive the random contribution, it is convenient to represent the dissipative force in Eq. (7) in a tensorial form as $\mathbf{F}_{ij}^D = -\boldsymbol{\mathcal{T}}_{ij}\cdot\mathbf{v}_{ij}$, where $\boldsymbol{\mathcal{T}}_{ij} = A(r_{ij})\mathbb{1} + B(r_{ij})\hat{\mathbf{e}}_{ij}\hat{\mathbf{e}}_{ij}$ and

$$A(r_{ij}) = \left(\frac{5\eta}{3} - \xi\right)\frac{F_{ij}}{\rho_i\rho_j} \quad \text{and} \quad B(r_{ij}) = 5\left(\frac{\eta}{3} + \xi\right)\frac{F_{ij}}{\rho_i\rho_j}. \tag{11}$$

Following the general framework of the fluid particle model[18], a random force can be defined as

$$\tilde{\mathbf{F}}_{ij}dt = \sqrt{2k_BT}\left(\tilde{A}(r_{ij})d\overline{\boldsymbol{\mathcal{W}}}^S_{ij} + \frac{\tilde{B}(r_{ij})}{3}\mathrm{tr}[d\boldsymbol{\mathcal{W}}_{ij}]\mathbb{1} + \tilde{C}(r_{ij})d\boldsymbol{\mathcal{W}}^A_{ij}\right)\cdot\hat{\mathbf{e}}_{ij}, \qquad (12)$$

where $d\boldsymbol{\mathcal{W}}_{ij}$ is a matrix of independent Wiener increments, $\mathrm{tr}[d\boldsymbol{\mathcal{W}}_{ij}]$ is the trace of this matrix, $d\overline{\boldsymbol{\mathcal{W}}}^S_{ij} = \frac{1}{2}(d\boldsymbol{\mathcal{W}}_{ij} + d\boldsymbol{\mathcal{W}}_{ji}) - \frac{1}{3}\mathrm{tr}[d\boldsymbol{\mathcal{W}}_{ij}]$ is the traceless symmetric part, and $d\boldsymbol{\mathcal{W}}^A_{ij} = \frac{1}{2}(d\boldsymbol{\mathcal{W}}_{ij} - d\boldsymbol{\mathcal{W}}_{ji})$ is the antisymmetric part. The functions $\tilde{A}(r)$, $\tilde{B}(r)$, and $\tilde{C}(r)$ are related to $A(r)$ and $B(r)$ of the tensor $\boldsymbol{\mathcal{T}}_{ij}$ as $A(r) = (\tilde{A}(r)^2 + \tilde{C}(r)^2)/2$ and $B(r) = (\tilde{A}(r)^2 - \tilde{C}(r)^2)/2 + (\tilde{B}(r)^2 - \tilde{A}(r)^2)/3$. To further simplify the expression in Eq. (12), we can select $\tilde{C}(r) = 0$ leading to

$$\tilde{A}(r_{ij}) = \left(2\left(\frac{5\eta}{3} - \xi\right)\frac{F_{ij}}{\rho_i\rho_j}\right)^{1/2} \quad \text{and} \quad \tilde{B}(r_{ij}) = \left(2\left(\frac{5\eta}{3} + 8\xi\right)\frac{F_{ij}}{\rho_i\rho_j}\right)^{1/2}. \quad (13)$$

The combination of forces in Eqs. (7) and (12) constitute the SDPD method, which describes hydrodynamics with consistent thermal fluctuations. The evolution of particle position and velocity follows Eq. (8). The derivation of SDPD in Ref.[9] also includes an energy conservation equation, which we omit here for simplicity. Note that this formulation violates the conservation of angular momentum, which can be essential in some physical problems (several examples will follow). The SDPD method without angular momentum conservation[9] will be called **SDPD-a**.

4 SDPD with Angular Momentum Conservation

In SDPD, dissipative and random forces possess not only a part along the inter-particle axis as in DPD, but also a component perpendicular to the inter-particle axis. This perpendicular part of dissipative and random forces destroys local and global angular momentum conservation. There exist a version of the SDPD method with angular momentum conservation[25], where the perpendicular component of dissipative and random forces has been neglected resulting in a method formulation very similar to DPD. In this method the input viscosity has to be scaled by a theoretically defined coefficient which depends on space dimension. However, it is advantageous to keep a perpendicular component of the dissipative force, since it provides much more efficient control over fluid transport properties than the component along inter-particle axis alone[26].

To extend the original SDPD formulation[9], we introduce a spin variable for every particle ω_i. In addition, each particle will also possess a moment of inertia I_i analogously to the already defined particle mass. In order to obtain discretized equations for the SDPD formulation with spin, we consider the NS equation with spin[27],

$$\rho\frac{d\mathbf{v}}{dt} = -\nabla p + (\eta + \eta_r)\nabla^2\mathbf{v} + \left(\frac{\eta}{3} + \xi - \eta_r\right)\nabla\nabla\cdot\mathbf{v} + 2\eta_r\nabla\times\omega, \qquad (14)$$

where η_r is the rotational viscosity and ω is the spin angular velocity. The introduced spin variable can be interpreted in two different ways. On the one hand, it is an approach used to recover angular momentum conservation in the SDPD formulation. On the other hand, the spin can be thought of as an effective angular velocity of a fluid volume represented by

a particle. However, it should not be confused with a molecular spin. The discretization of the NS equation with spin provides a consistent model, where translational and rotational friction interactions are properly balanced unlike the FPM model which does not have a direct connection to the NS equation.

Following the rules for derivatives in Appendix A we obtain the three forces: conservative (C), dissipative (D), and rotational (R) given by

$$\begin{aligned}
\mathbf{F}^C_{ij} &= \left(\frac{p_i}{\rho_i^2}+\frac{p_j}{\rho_j^2}\right)F_{ij}\mathbf{r}_{ij},\\
\mathbf{F}^D_{ij} &= -\left(\frac{5\eta}{3}+3\eta_r-\xi\right)\frac{F_{ij}}{\rho_i\rho_j}\mathbf{v}_{ij}\ -5\left(\frac{\eta}{3}+\xi-\eta_r\right)\frac{F_{ij}}{\rho_i\rho_j}\hat{\mathbf{e}}_{ij}(\hat{\mathbf{e}}_{ij}\cdot\mathbf{v}_{ij}),\\
\mathbf{F}^R_{ij} &= -2\eta_r\frac{F_{ij}}{\rho_i\rho_j}\mathbf{r}_{ij}\times(\boldsymbol{\omega}_i+\boldsymbol{\omega}_j)\,.
\end{aligned} \tag{15}$$

The conservative and dissipative forces are very similar to the expressions in Eq. (7). In addition, now we have a rotational force which is also dissipative, but acts on particles' angular velocities such that a spin of one particle leads to a change in translational and angular velocity of another particle.

The defined set of deterministic forces in Eq. (15) can be referred to as a SPH discretization with angular momentum conservation. However, the SDPD method also incorporates consistently thermal fluctuations by appending a random force to the set of forces in Eq. (15). Here, the combination of dissipative, rotational, and random forces has to satisfy the fluctuation-dissipation balance. Similarly to the dissipative force, we re-write the rotational force in a tensorial form as $\mathbf{F}^R_{ij}=-\boldsymbol{\mathcal{T}}_{ij}\cdot(\mathbf{r}_{ij}\times(\boldsymbol{\omega}_i+\boldsymbol{\omega}_j))/2$. Note that the same tensor $\boldsymbol{\mathcal{T}}_{ij}$ is used for both dissipative and rotational forces in order to satisfy the fluctuation-dissipation theorem when a random force is added[18]. This implies that $\eta_r=5\eta/3-\xi$ and

$$A(r_{ij})=4\left(\frac{5\eta}{3}-\xi\right)\frac{F_{ij}}{\rho_i\rho_j},\quad\text{and}\quad B(r_{ij})=10\left(\xi-\frac{2\eta}{3}\right)\frac{F_{ij}}{\rho_i\rho_j}. \tag{16}$$

Then, the corresponding functions for the random force in Eq. (12) are given by

$$\tilde{A}(r_{ij})=\left(2\left(\frac{20\eta}{3}-4\xi\right)\frac{F_{ij}}{r_ir_j}\right)^{1/2}\quad\text{and}\quad\tilde{B}(r_{ij})=\left(2\left(17\xi-\frac{40\eta}{3}\right)\frac{F_{ij}}{r_ir_j}\right)^{1/2}. \tag{17}$$

The full set of forces for the SDPD method with angular momentum conservation is finally given by

$$\begin{aligned}
\mathbf{F}^C_{ij} &= \left(\frac{p_i}{\rho_i^2}+\frac{p_j}{\rho_j^2}\right)F_{ij}\mathbf{r}_{ij},\\
\mathbf{F}^D_{ij} &= -\left(\gamma^a_{ij}\left(\mathbf{v}_{ij}+\frac{\hat{\mathbf{e}}_{ij}\,(\hat{\mathbf{e}}_{ij}\cdot\mathbf{v}_{ij})}{3}\right)\right)-\frac{2\gamma^b_{ij}}{3}\hat{\mathbf{e}}_{ij}\,(\hat{\mathbf{e}}_{ij}\cdot\mathbf{v}_{ij})\,,\\
\mathbf{F}^R_{ij} &= -\gamma^a_{ij}\frac{\mathbf{r}_{ij}}{2}\times(\boldsymbol{\omega}_i+\boldsymbol{\omega}_j)\,,\\
\tilde{\mathbf{F}}_{ij} &= \left(\sigma^a_{ij}d\overline{\boldsymbol{\mathcal{W}}}_{ij}+\sigma^b_{ij}\frac{1}{3}tr[d\boldsymbol{\mathcal{W}}_{ij}]\mathbb{1}\right)\cdot\frac{\hat{\mathbf{e}}_{ij}}{dt},
\end{aligned} \tag{18}$$

where

$$\gamma_{ij}^a = \left(\frac{20\eta}{3} - 4\xi\right)\frac{F_{ij}}{\rho_i\rho_j}, \quad \gamma_{ij}^b = \left(17\xi - \frac{40\eta}{3}\right)\frac{F_{ij}}{\rho_i\rho_j}, \tag{19}$$

and $\sigma_{ij}^{a,b} = 2\sqrt{k_B T \gamma_{ij}^{a,b}}$. It is important to note that these equations are only valid for $2\eta/3 \leq \xi \leq 5\eta/3$, such that the friction coefficients $(\gamma_{ij}^a + 2\gamma_{ij}^b)/3$ and γ_{ij}^a are positive. Another simplification which can be made is the reduction to a single dissipative parameter γ_{ij} such that $\xi = 20\eta/21$ and

$$\gamma_{ij}^a = \gamma_{ij}^b = \gamma_{ij} = \frac{20\eta}{7}\frac{F_{ij}}{\rho_i\rho_j}, \qquad \sigma_{ij}^a = \sigma_{ij}^b = \sigma_{ij} = 2\sqrt{k_B T \gamma_{ij}}. \tag{20}$$

The derived SDPD method with angular momentum conservation will be referred to as **SDPD+a** further in text. To describe time evolution of the position and of the translational and angular velocity of a particle i, Eq. (8) with an addition of angular velocity integration ($\dot{\boldsymbol{\omega}}_i = \sum_j \mathbf{N}_{ij}/I_j$) is employed, where $\mathbf{N}_{ij}$ is the torque exerted by particle j on particle i and is given by $\mathbf{N}_{ij} = \mathbf{r}_{ij} \times \mathbf{F}_{ij}/2$. This leads to local and global angular momentum conservation. More details on the SDPD method with angular momentum conservation can be found in Ref.[17].

5 Application of the SDPD Method

The SDPD method has been used in simulations of colloidal particles[11] and polymers[12], single red blood cells in tube flow[13], margination of leukocytes[14], and margination of micro- and nano-particles[15] in blood flow.

5.1 Margination of micro- and nano-particles in blood flow

The first step in the delivery of small drug carriers is their transport along blood vessels which determines their initial distribution. Further, the distribution of micro- and nano-particles is affected by their binding to specific targeted sites. However, efficient binding of carriers can be achieved only if they are present near vessel walls at sufficiently high concentrations, and thus, the distribution of micro- and nano-carriers within vessel cross-sections plays an essential role in their efficient delivery. The cross-sectional distribution of micro- and nano-particles depends on local blood flow properties such as hematocrit (volume fraction of red blood cells) and flow rate as well as on the particle characteristics such as their size, shape, and deformability. The migration of various suspended particles toward walls in blood flow, which is also often referred to as *margination*, has been observed for white blood cells[28,29,14], platelets[30,31], and rigid micro-particles[32,33]. Particle margination in blood flow is mediated by red blood cells (RBCs), which migrate to the vessel center[34] due to hydrodynamic interactions with the walls (called lift force)[35–37], and lead to an increased concentration of different particles within the RBC-free layer (RBCFL) near a wall (a layer void of RBCs). More precisely, the margination mechanism is a consequence of the competition between lift forces on RBCs and suspended particles, and their interactions in flow[38]. Similarly, micro- and nano-carriers have a potential to get marginated[15], and therefore to interact with vessel walls.

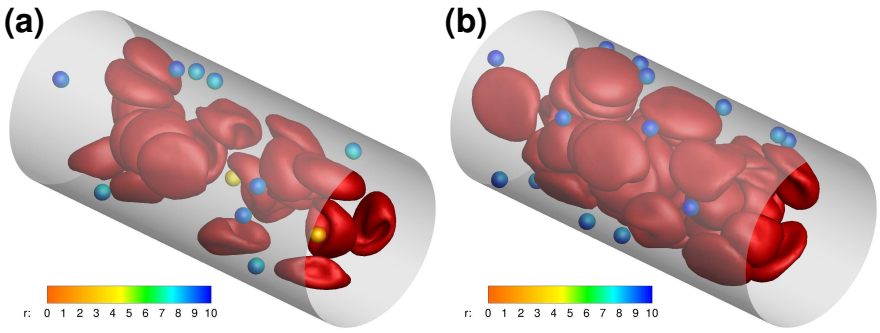


Figure 1: Particle distributions in blood flow. Illustrations of 3D simulations of blood flow for different hematocrit values (a) $H_t = 0.2$ and (b) $H_t = 0.4$. RBCs are drawn in red, while spherical carriers with a size of $D_p = 1.83\,\mu\text{m}$ are colored according to their radial position r. For better contrast, carrier positions from several time instances are superimposed in the plots.

Numerical simulations of blood flow on a single-cell level allow us to explore the flow behavior and interaction of blood cells and other suspended components[39–41]. Thus, the role of particle size and shape on the margination efficiency can be investigated *in silico*[15] providing information about particle adhesion potential, since particle margination is an essential pre-condition for particle adhesion. The findings in Ref.[15] indicate that spherical particles have slightly better margination properties than ellipsoids, however the adhesion efficiency of ellipsoidal particles appears to be superior in comparison with that of spheres due to a larger area for adhesive interactions[42]. The larger particles possess a larger probability of being marginated. As the particle size becomes very small (less than about $100 - 200$ nm), the particle distribution within vessel cross-section can be described well by the volume of blood plasma, since small particles are uniformly distributed within the suspending fluid.

The simulations are based on the SDPD-a method. Blood is modeled as a suspension of RBCs and micro- or nano-particles. RBCs and suspended micro- and nano-particles are modeled by a collection of particles on their surface connected by viscoelastic springs. The membrane model also incorporates bending resistance and the area and volume conservation constraints. For full details about the membrane model and simulation conditions and setup we refer the reader to Ref.[15].

Margination of micro- and nano-particles in blood flow depends on hematocrit H_t, vessel diameter, and flow rate[15]. Figure 1 illustrates the distribution of carriers with a diameter $D_p = 1.83\,\mu\text{m}$ for two H_t values. For better visibility, the carrier positions from a few snapshots are superimposed in the plot. The carrier surfaces are colored according to their radial position in the channel, with yellow color indicating a position near the channel center, while blue color corresponds to a position near the wall. Clearly, the carriers are marginating better for the case of larger H_t.

Carrier positions in blood flow sampled over time lead to particle distributions, which reflect the probability of a particle to be at a certain distance from the wall. Figure 2 shows several center-of-mass distributions of circular particles in 2D with $D_p = 1.83\,\mu\text{m}$ for several H_t values. The RBC-free layer (RBCFL) thickness, which is computed from simulation snapshots through the analysis of the RBC core boundary[43] similar to experimental

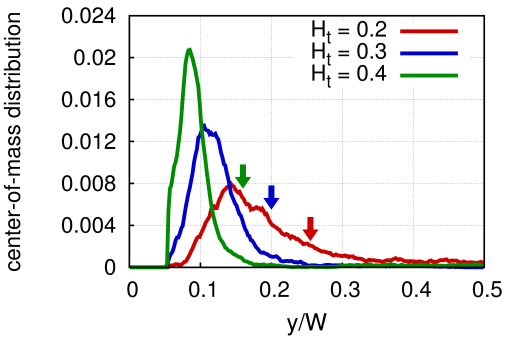


Figure 2: Center-of-mass distributions of carriers for various H_t values. 2D simulation results for circular particles with $D_p = 1.83\,\mu\text{m}$. The wall is at $y/W = 0$, where W is the channel width. The arrows indicate the boundary of the RBCFL for the different hematocrits, marked by corresponding colors.

measurements[44], is depicted by small arrows. The distributions have been averaged over the halves of the channel due to symmetry. Figure 2 shows that the carriers migrate into the RBCFL and remain quasi-trapped there. With increasing H_t, the carriers marginate better, as indicated by the development of a strong peak in the distribution near the wall at $y/W = 0$ (W is the channel width), and the motion of the peak position towards the wall. This is due to a decrease in the RBCFL thickness leading to a smaller available space for the particles. This trend is in agreement with experimental observations[32] and simulations[45–47] of margination of blood platelets, which have a comparable size. Further details on micro- and nano-particle margination in blood flow depending on various conditions can be found in Ref.[15].

5.2 Taylor-Couette flow of two immiscible fluids

To illustrate the importance of angular momentum conservation, we test the SDPD method on Taylor-Couette flow, which usually refers to a fluid flow in the gap between two rotating cylinders as shown in Fig. 3. However, we consider a setup, where the inner cylinder is replaced by another immiscible fluid such that no mixing between the two fluids at R_i can occur. A solution of the incompressible NS equation for this problem yields a linear angular velocity profile $v_\phi(r) = \Omega_0 r$ across both immiscible fluids, where r is the radial position. Note that this solution is independent of the viscosity values of the immiscible fluids.

Recent numerical simulations with a similar setup[16] have shown that the conservation of angular momentum is necessary to obtain correct velocity and torque profiles across immiscible fluids with different viscosities. In simulations, the ratio of fluid viscosities was set to $\eta_i/\eta_o = 3$. The computational domain was assumed to be periodic in the cylinder-axis direction, while the cylinder wall was modeled by a layer of frozen particles with a thickness h whose structure (e.g., radial distribution function) was the same as that of the fluids. To prevent mixing of the fluids and particle penetration into the wall, specular reflection of particles has been imposed at cylindrical surfaces with $r = R_i$ and

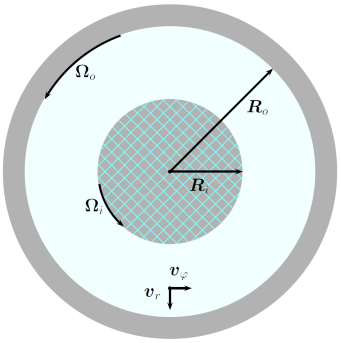


Figure 3: A sketch of two cylinders with radii R_o and R_i where the subscripts 'o' and 'i' denote the outer and inner cylinders, respectively. The gap between cylinders is filled with a fluid colored in blue. Taylor-Couette flow can be generated by rotation of the outer cylinder with a rotational frequency Ω_o. In simulations with two immiscible fluids, the inner cylinder (shaded area) is replaced by another fluid which cannot mix with the fluid inside the gap between two cylindrical surfaces.

$r = R_o$. The wall particles were rotated with a constant angular frequency Ω_o in order to generate flow. Figure 4 shows angular velocity profiles for the Taylor-Couette flow using both SDPD+a and SDPD-a methods. The SDPD+a simulation properly captures a linear profile of angular velocity, while the SDPD-a method leads to distinct slopes within the regions of different viscosities.

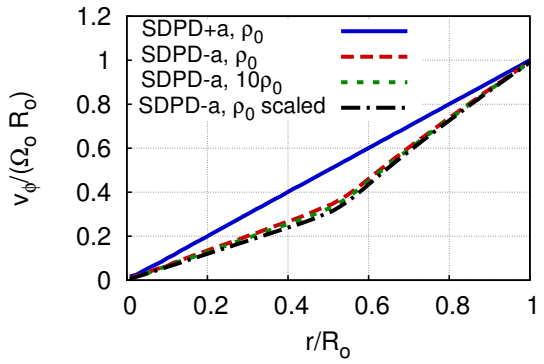


Figure 4: Angular velocity v_ϕ profiles for the Taylor-Couette flow with two immiscible fluids using both SDPD+a and SDPD-a methods. Radial position r is normalized by the cylinder radius R_o, while angular velocity is scaled with the cylinder angular velocity $\Omega_o R_o$. The SDPD+a method leads to a correct linear profile, while the SDPD-a method fails to do so due to violation of angular momentum conservation. The SDPD-a results for different resolution with $\rho = 10\rho_0$ and for a twice larger system size (marked as "scaled") show hardly any dependence on fluid resolution.

6 Tank-Treading of a Vesicle in Shear Flow

Flow dynamics of soft deformable objects such as liquid droplets, lipid vesicles, red blood cells, and elastic capsules has attracted a lot of scientific interest recently due to a wide range of possible applications. For instance, a number of experiments[36,48], theoretical approaches[21,49,50], and simulations[50–52,39] have shown that fluid vesicles exhibit a rich dynamical behavior in shear flow including tank-treading (TT) and tumbling (TB) motion. The tumbling motion corresponds to vesicle rotation around its center-of-mass nearly as a rigid body. A tank-treading vesicle in shear flow shows a stationary shape with a finite inclination angle $\theta > 0$ with respect to the flow direction, while the membrane is rotating around the center-of-mass of the vesicle, see Fig. 5. The occurrence of different vesicle motion is governed by the viscosity contrast $\lambda = \eta_i/\eta_o$ between fluids inside and outside the vesicle with viscosities η_i and η_o, respectively. A physical explanation for the TT-to-TB transition can be derived from the two components of shear flow: an elongational part which tends to stretch and align a vesicle along the $x = y$ axis with an inclination angle of $\theta = \pi/4$ and a rotational part of the flow which exerts a torque on the vesicle membrane. Increasing of viscosity contrast leads to higher shear stresses inside the vesicle opposing its TT motion, which results in an effective torque and decrease of the vesicle inclination angle. Thus, for high enough λ a transition from TT to TB motion occurs. Keller and Skalak (KS)[21] derived a theory which predicts the TT-to-TB transition. Moreover, the KS theory is able to predict the inclination angle θ in the vesicle TT regime. Details of the KS theory are given in Appendix B.

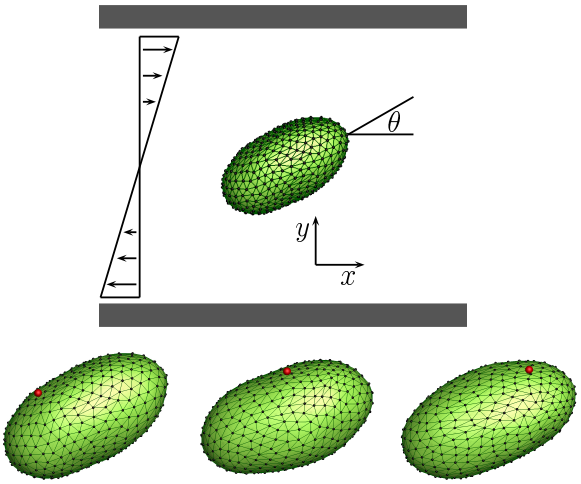


Figure 5: Simulation snapshots of a tank-treading vesicle in shear flow for $\lambda = 2$. A xy-plane view, where the flow is in x direction. The red sphere is attached to a fixed position on a vesicle in order to illustrate the TT motion of the membrane; however, it is just a marker used for visualization and introduced at post-processing stage. Note that small shape fluctuations are clearly visible.

The vesicle membrane is modeled by a collection of particles on an ellipsoidal surface. An illustration of a vesicle structure is shown in Fig. 5. The model incorporates membrane

bending rigidity, the constraints for vesicle area and volume, and the viscosity contrast between inner and outer fluids; more details on vesicle modeling can be found in Ref.[17]. The simulated ellipsoidal vesicle has a prolate shape with $a_1 > a_2 = a_3$ and an aspect ratio of $a_1/a_2 \approx 1.7$. The vesicle is placed in a box, where periodic boundary conditions are applied in x and z direction, while shear flow is generated in the x direction, as shown in Fig. 5.

The inclination angle of a TT vesicle in shear flow is calculated by

$$\theta = \arctan(u_y/u_x), \tag{21}$$

where $\mathbf{u} = (u_x, u_y, u_z)$ is the eigenvector of the moments of inertia tensor with the smallest eigenvalue. Figure 6 compares inclination angles obtained from simulations with SDPD+a and SDPD-a fluids and from the KS theory for different viscosity ratios λ. The simulation results obtained with a SDPD+a fluid agree very well with the KS theory predictions, while the results using a SDPD-a fluid show a significant overestimation of the inclination angle at large λ. The results for $\lambda = 1$ from both SDPD+a and SDPD-a cases coincide indicating that angular momentum conservation does not affect simulation results if inner and outer fluids have the same viscosity. The deviations of the SDPD+a results from the KS theory predictions might be due to small shape fluctuations of the vesicle and/or numerical errors indicated by error bars in Fig. 6.

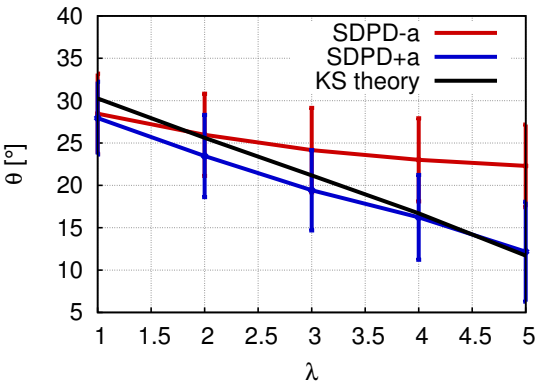


Figure 6: Comparison of inclination angles θ of a TT vesicle in shear flow obtained from SDPD+a (blue) and SDPD-a (red) simulations and from the KS theory (black) for different viscosity ratios λ.

7 Summary

We presented an overview of the SDPD-a method[9] and its extension[17], which satisfies local and global angular momentum conservation. In the SDPD+a, each particle possesses an angular velocity, and its rotational contribution has been derived from the NS equation with spin following the SPH formalism. This leads to a spin variable similar to the FPM model[18]. Thermal fluctuations in SDPD+a have been also included similar to those in FPM

[18]. Several applications of the SDPD method have been considered. The first example concerned the margination of micro- and nano-particles in blood flow, which is important in drug delivery. Simulations of Taylor-Couette flow with two immiscible fluids show that SDPD+a leads to correct predictions of flow profiles in agreement with analytical results, while SDPD-a fails to capture properly flow characteristics due to violation of angular momentum conservation. Finally, simulations of vesicle dynamics in shear flow reveal that angular momentum conservation is essential to obtain correct results for the inclination angle of a tank-treading vesicle if there exists a viscosity contrast λ between inner and outer fluids. For $\lambda \neq 1$ the SDPD+a method predicts vesicle characteristics in agreement with the KS theory for a vesicle in shear flow, while SDPD-a overestimates the inclination angle. In conclusion, the family of SDPD methods including the version with angular momentum conservation provides valuable tools for modeling flows of complex fluids.

Acknowledgments

This work has been supported by the DFG Research Unit FOR 1543 SHENC – Shear Flow Regulation in Hemostasis. Dmitry A. Fedosov acknowledges funding by the Alexander von Humboldt Foundation. Kathrin Müller acknowledges support by the International Helmholtz Research School of Biophysics and Soft Matter (IHRS BioSoft). We also gratefully acknowledge a CPU time grant by the Jülich Supercomputing Center.

Appendix

A Calculation of Derivatives

We summarize the calculation of derivatives of field variables similar to those in SPH[22]. Using Eq. (4), the first derivative of a field g can be approximated as

$$\frac{\partial \tilde{g}}{\partial x} = \sum_{j=1}^{N} \frac{m_j}{\rho_j} g \frac{\partial W_{ij}}{\partial x_i}, \tag{22}$$

where the notations are identical to those in the main text. A disadvantage of this approximation is that the derivative does not vanish for g being a constant function. Therefore, a better approximation is given by

$$\frac{\partial \tilde{g}}{\partial x} = \frac{1}{\phi} \left(\frac{\partial (\tilde{g}\phi)}{\partial x} - \tilde{g} \frac{\partial \phi}{\partial x} \right), \tag{23}$$

where ϕ must be a differentiable function. Following Eq. (22), we then obtain

$$\frac{\partial \tilde{g}}{\partial x} = \frac{1}{\phi_i} \sum_{j=1}^{N} \frac{m_j}{\rho_j} \phi_j \left(g_j - g_i \right) \frac{\partial W_{ij}}{\partial x_i}. \tag{24}$$

When $\phi = 1$, Eq. (24) reduces to

$$\frac{\partial \tilde{g}}{\partial x} \approx \sum_{j} \frac{m_j}{\rho_j} g_{ji} \frac{\partial W_{ij}}{\partial x_i}, \tag{25}$$

where $g_{ji} = g_j - g_i$. In Eq. (24), $\phi = \rho$ can be also selected, yielding an approximation for the first derivative as

$$\frac{\partial \tilde{g}}{\partial x} \approx \frac{1}{\rho_i} \sum_j m_j g_{ji} \frac{\partial W_{ij}}{\partial x_i}. \tag{26}$$

The choice for different discretizations ($\phi = 1$ or $\phi = \rho$) may depend on a problem of interest. For instance, when different interacting fluids with large density ratios are considered, it has been shown that the approximation in Eq. (24) with $\phi = 1$ is more accurate than that with $\phi = \rho$, because ρ in Eq. (25) is included directly inside the sum[53,22]. Furthermore, if only a single fluid is employed, an approximation $\phi_i \approx \rho_j$ can be used making the above choices for ϕ equivalent.

There exists another definition for the first derivative,

$$\frac{\partial \tilde{g}}{\partial x} = \phi \left(\frac{\partial}{\partial x} \left(\frac{\tilde{g}}{\phi} \right) + \frac{\tilde{g}}{\phi^2} \frac{\partial \phi}{\partial x} \right). \tag{27}$$

Following the SPH formalism[22] we obtain

$$\frac{\partial \tilde{g}}{\partial x} \approx \phi_i \sum_{j=1}^{N} \frac{m_j}{\rho_j} \left(\frac{g_j}{\phi_j} + \frac{g_i}{\phi_i^2} \phi_j \right) \frac{\partial W_{ij}}{\partial x_i}. \tag{28}$$

As a result, a choice of $\phi = 1$ here leads to

$$\frac{\partial \tilde{g}}{\partial x} \approx \sum_{j=1}^{N} \frac{m_j}{\rho_j} \left(g_j + g_i \right) \frac{\partial W_{ij}}{\partial x_i}, \tag{29}$$

while for $\phi \approx \rho$ Eq. (28) becomes

$$\frac{\partial \tilde{g}}{\partial x} \approx \rho_i \sum_{j=1}^{N} m_j \left(\frac{g_j}{\rho_j^2} + \frac{g_i}{\rho_i^2} \right) \frac{\partial W_{ij}}{\partial x_i}. \tag{30}$$

A set of equations above defines different approximations of first derivatives, which can be used to derive discretizations of other differential operators in the NS equation. For instance, using Eq. (26) the gradient of $g(\mathbf{r})$ can be approximated as

$$\nabla \tilde{g} \approx -\frac{1}{\rho_i} \sum_j m_j g_{ij} \nabla_i W_{ij}, \tag{31}$$

where $g_{ij} = g_i - g_j$. Similarly, the divergence and the curl of a vector field $\tilde{\mathbf{G}}(\mathbf{r})$ are discretized as

$$\nabla_i \cdot \tilde{\mathbf{G}}_i \approx -\frac{1}{\rho_i} \sum_j m_j \mathbf{G}_{ij} \cdot \nabla_i W_{ij}, \tag{32}$$

$$\nabla_i \times \tilde{\mathbf{G}}_i \approx -\rho_i \sum_j \frac{m_j}{\rho_i \rho_j} \left(\mathbf{G}_j + \mathbf{G}_i \right) \times \nabla_i W_{ij}. \tag{33}$$

The second derivatives are then given by

$$\nabla_i \left(\nabla_i \cdot \tilde{\mathbf{G}}_i \right) \approx -\sum_j m_j \frac{F_{ij}}{\rho_i \rho_j} \left(5 \hat{\mathbf{e}}_{ij} \left(\hat{\mathbf{e}}_{ij} \cdot \mathbf{G}_{ij} \right) - \mathbf{G}_{ij} \right) \tag{34}$$

and

$$\nabla_i^2 \tilde{g}_i \approx -2 \sum_j m_j \frac{F_{ij}}{\rho_i \rho_j} g_{ij}, \tag{35}$$

where $\hat{\mathbf{e}}_{ij} = \mathbf{r}_{ij}/r_{ij}$ is the unity vector along the separation direction of particles i and j[9].

The curl of a vector field $\mathbf{G}$ can be approximated as

$$\nabla_i \times \tilde{\mathbf{G}}_i \approx \phi_i \sum_j \frac{m_j}{\rho_j} \nabla_i W(r_{ij}) \times \left(\frac{\mathbf{G}_j}{\phi_j} + \frac{\phi_j \mathbf{G}_i}{\phi_i^2} \right), \tag{36}$$

where a selection of $\phi = 1$ leads to

$$\nabla_i \times \tilde{\mathbf{G}}_i \approx \sum_j \frac{m_j}{\rho_j} \nabla_i W(r_{ij}) \times (\mathbf{G}_j + \mathbf{G}_i), \tag{37}$$

while $\phi = \rho$ results in

$$\nabla_i \times \tilde{\mathbf{G}}_i \approx \rho_i \sum_j m_j \nabla_i W(r_{ij}) \times \left(\frac{\mathbf{G}_j}{\rho_j^2} + \frac{\mathbf{G}_i}{\rho_i^2} \right). \tag{38}$$

B Keller-Skalak Theory

The Keller-Skalak (KS) theory[21] assumes a fixed ellipsoidal shape $(r_1/a_1)^2 + (r_2/a_2)^2 + (r_3/a_3)^2 = 1$, where $r_i, i \in \{1, 2, 3\}$ are the Cartesian coordinates and a_i are the semiaxes of the ellipsoid. The motion of a vesicle is derived by considering energy balance between the energy supplied by the fluid and the energy which dissipates on the membrane and inside the vesicle. This balance leads to a differential equation given by

$$\frac{d\theta}{dt} = \frac{1}{2} \dot{\gamma} \left(B \cos(2\theta) - 1 \right), \tag{39}$$

where $\dot{\gamma}$ is the shear rate. If $B > 1$, the vesicle is in the TT regime, and hence, a steady inclination angle can be found as $\theta = 0.5 \arccos(1/B)$, where B is a function of vesicle shape and viscosity contrast given by

$$\begin{aligned}
B &= f_0 \left(f_1 + \frac{1}{f_1} \left(\frac{1}{1 + f_2(\lambda - 1)} \right) \right), \\
f_0 &= \frac{2}{a_1/a_2 + a2_/a_1}, \\
f_1 &= 0.5 \left(a_1/a_2 - a_2/a_1 \right), \\
f_2 &= 0.5 g(\alpha_1^2 + \alpha_2^2), \\
g &= \int_0^\infty (\alpha_1^2 + s)^{-3/2} (\alpha_2^2 + s)^{-3/2} (\alpha_3^2 + s)^{-1/2} \, ds, \\
\alpha_i &= \frac{a_i}{a_1 a_2 a_3}.
\end{aligned} \tag{40}$$

Note that the KS theory does not consider vesicle's membrane viscosity.

References

1. S. Succi, *The Lattice Boltzmann equation for fluid dynamics and beyond*, Oxford University Press, Oxford, 2001.
2. P. J. Hoogerbrugge and J. M. V. A. Koelman, *Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics*, Europhys. Lett., **19**, 155–160, 1992.
3. P. Español and P. Warren, *Statistical mechanics of dissipative particle dynamics*, Europhys. Lett., **30**, 191–196, 1995.
4. R. D. Groot and P. B. Warren, *Dissipative particle dynamics: bridging the gap between atomistic and mesoscopic simulation*, J. Chem. Phys., **107**, 4423–4435, 1997.
5. A. Malevanets and R. Kapral, *Mesoscopic model for solvent dynamics*, J. Chem. Phys., **110**, 8605–8613, 1999.
6. G. Gompper, T. Ihle, D. M. Kroll, and R. G. Winkler, *Multi-particle collision dynamics: a particle-based mesoscale simulation approach to the hydrodynamics of complex fluids*, Adv. Polym. Sci., **221**, 1–87, 2009.
7. L. B. Lucy, *A numerical approach to the testing the fission hypothesis*, Astronom. J., **82**, 1013–1024, 1977.
8. J. J. Monaghan, *Smoothed particle hydrodynamics*, Annu. Rev. Astron. Astrophys., **30**, 543–574, 1992.
9. P. Español and M. Revenga, *Smoothed dissipative particle dynamics*, Phys. Rev. E, **67**, 026705, 2003.
10. A. Vázquez-Quesada, M. Ellero, and P. Español, *Consistent scaling of thermal fluctuations in smoothed dissipative particle dynamics*, J. Chem. Phys., **130**, 034901, 2009.
11. X. Bian, S. Litvinov, R. Qian, M. Ellero, and N. A. Adams, *Multiscale modeling of particle in suspension with smoothed dissipative particle dynamics*, Phys. Fluids, **24**, 012002, 2012.
12. S. Litvinov, M. Ellero, X. Hu, and N. A. Adams, *Smoothed dissipative particle dynamics model for polymer molecules in suspension*, Phys. Rev. E, **77**, 066703, 2008.
13. D. A. Fedosov, M. Peltomäki, and G. Gompper, *Deformation and dynamics of red blood cells in flow through cylindrical microchannels*, Soft Matter, **10**, 4258–4267, 2014.
14. D. A. Fedosov and G. Gompper, *White blood cell margination in microcirculation*, Soft Matter, **10**, 2961–2970, 2014.
15. K. Müller, D. A. Fedosov, and G. Gompper, *Margination of micro- and nano-particles in blood flow and its effect on drug delivery*, Sci. Rep., **4**, 4871, 2014.
16. I. O. Götze, H. Noguchi, and G. Gompper, *Relevance of angular momentum conservation in mesoscale hydrodynamics simulations*, Phys. Rev. E, **76**, 046705, 2007.
17. K. Müller, D. A. Fedosov, and G. Gompper, *Smoothed dissipative particle dynamics with angular momentum conservation*, J. Comp. Phys., **281**, 301–315, 2015.
18. P. Español, *Fluid particle model*, Phys. Rev. E, **57**, 2930–2948, 1998.
19. W. Pan, I. V. Pivkin, and G. E. Karniadakis, *Single-particle hydrodynamics in DPD: a new formulation*, Europhys. Lett., **84**, 10012, 2008.
20. S. R. de Groot and P. Mazur, *Non-equilibrium thermodynamics*, North-Holland, Amsterdam, 1962.
21. S. R. Keller and R. Skalak, *Motion of a tank-treading ellipsoidal particle in a shear*

flow, J. Fluid Mech., **120**, 27–47, 1982.
22. J. J. Monaghan, *Smoothed particle hydrodynamics*, Rep. Prog. Phys., **68**, 1703–1759, 2005.
23. M. P. Allen and D. J. Tildesley, *Computer simulation of liquids*, Clarendon Press, New York, 1991.
24. G. K. Batchelor, *An introduction to fluid dynamics*, Cambridge University Press, Cambridge, 2000.
25. X. Y. Hu and N. A. Adams, *Angular-momentum conservative smoothed particle dynamics for incompressible viscous flows*, Phys. Fluids, **18**, 101702, 2006.
26. C. Junghans, M. Praprotnik, and K. Kremer, *Transport properties controlled by a thermostat: an extended dissipative particle dynamics thermostat*, Soft Matter, **4**, 156–161, 2008.
27. D. W. Condiff and J. S. Dahler, *Fluid mechanical aspects of antisymmetric stress*, Phys. Fluids, **7**, 842–854, 1964.
28. U. Bagge and R. Karlsson, *Maintenance of white blood cell margination at the passage through small venular junctions*, Microvasc. Res., **20**, 92–95, 1980.
29. H. L. Goldsmith and S. Spain, *Margination of leukocytes in blood flow through small tubes*, Microvasc. Res., **27**, 204–222, 1984.
30. G. J. Tangelder, H. C. Teirlinck, D. W. Slaaf, and R. S. Reneman, *Distribution of blood platelets flowing in arterioles*, Am. J. Physiol., **248**, H318–H323, 1985.
31. B. Woldhuis, G. J. Tangelder, D. W. Slaaf, and R. S. Reneman, *Concentration profile of blood platelets differs in arterioles and venules*, Am. J. Physiol., **262**, H1217–H1223, 1992.
32. A. W. Tilles and E. C. Eckstein, *The near-wall excess of platelet-sized particles in blood flow: its dependence on hematocrit and wall shear rate*, Microvasc. Res., **33**, 211–223, 1987.
33. E. C. Eckstein, A. W. Tilles, and F. J. Millero III, *Conditions for the occurrence of large near-wall excesses of small particles during blood flow*, Microvasc. Res., **36**, 31–39, 1988.
34. H. L. Goldsmith, G. R. Cokelet, and P. Gaehtgens, *Robin Fahraeus: evolution of his concepts in cardiovascular physiology*, Am. J. Physiol., **257**, H1005–H1015, 1989.
35. I. Cantat and C. Misbah, *Lift force and dynamical unbinding of adhering vesicles under shear flow*, Phys. Rev. Lett., **83**, 880–883, 1999.
36. M. Abkarian, C. Lartigue, and A. Viallat, *Tank treading and unbinding of deformable vesicles in shear flow: determination of the lift force*, Phys. Rev. Lett., **88**, 068103, 2002.
37. S. Messlinger, B. Schmidt, H. Noguchi, and G. Gompper, *Dynamical regimes and hydrodynamic lift of viscous vesicles under shear*, Phys. Rev. E, **80**, 011901, 2009.
38. A. Kumar and M. D. Graham, *Mechanism of margination in confined flows of blood and other multicomponent suspensions*, Phys. Rev. Lett., **109**, 108102, 2012.
39. D. A. Fedosov, H. Noguchi, and G. Gompper, *Multiscale modeling of blood flow: from single cells to blood rheology*, Biomech. Model. Mechanobiol., **13**, 239–258, 2014.
40. J. B. Freund, *Numerical simulation of flowing blood cells*, Annu. Rev. Fluid Mech., **46**, 67–95, 2014.
41. D. A. Fedosov, M. Dao, G. E. Karniadakis, and S. Suresh, *Computational biorheology*

of human blood flow in health and disease, Ann. Biomed. Eng., **42**, 368–387, 2014.
42. P. Decuzzi and M. Ferrari, *The adhesive strength of non-spherical particles mediated by specific interactions*, Biomaterials, **27**, 5307–5314, 2006.
43. D. A. Fedosov, B. Caswell, A. S. Popel, and G. E. Karniadakis, *Blood flow and cell-free layer in microvessels*, Microcirculation, **17**, 615–628, 2010.
44. S. Kim, R. L. Kong, A. S. Popel, M. Intaglietta, and P. C. Johnson, *Temporal and spatial variations of cell-free layer width in arterioles*, Am. J. Physiol., **293**, H1526–H1535, 2007.
45. L. Crowl and A. L. Fogelson, *Analysis of mechanisms for platelet near-wall excess under arterial blood flow conditions*, J. Fluid Mech., **676**, 348–375, 2011.
46. H. Zhao and E. S. G. Shaqfeh, *Shear-induced platelet margination in a microchannel*, Phys. Rev. E, **83**, 061924, 2011.
47. D. A. Reasor Jr, M. Mehrabadi, D. N. Ku, and C. K. Aidun, *Determination of critical parameters in platelet margination*, Ann. Biomed. Eng., **41**, 238–249, 2013.
48. V. Kantsler and V. Steinberg, *Transition to tumbling and two regimes of tumbling motion of a vesicle in shear flow*, Phys. Rev. Lett., **96**, 036001, 2006.
49. R. Tran-Son-Tay, S. P. Sutera, and P. R. Rao, *Determination of red blood cell membrane viscosity from rheoscopic observations of tank-treading motion*, Biophys. J., **46**, 65–72, 1984.
50. H. Noguchi and G. Gompper, *Dynamics of fluid vesicles in shear flow: effect of the membrane viscosity and thermal fluctuations*, Phys. Rev. E, **72**, 011901, 2005.
51. A. Yazdani and P. Bagchi, *Three-dimensional numerical simulation of vesicle dynamics using a front-tracking method*, Phys. Rev. E, **85**, 056308, 2012.
52. H. Zhao and E. S. G. Shaqfeh, *The dynamics of a vesicle in simple shear flow*, J. Fluid Mech., **674**, 578–604, 2011.
53. A. Colagrossi and M. Landrini, *Numerical simulation of interfacial flows by smoothed particle hydrodynamics*, J. Comp. Phys., **191**, 448–475, 2003.

Lattice-Boltzmann Simulations of Colloids and Polyelectrolytes in Electrolyte Solutions

Christian Holm and Owen A. Hickey

Institut für Computerphysik, Universität Stuttgart, Allmandring 3, 70569 Stuttgart, Germany

E-mail: holm@icp.uni-stuttgart.de

Electric fields are often employed in experiments in order to probe the properties of both pure electrolyte solutions and electrolyte solutions containing charged colloids and/or polyelectrolytes. These systems pose a problem for computer simulations as their properties are a result of coupled electrostatic and hydrodynamic interactions. Both of these interactions are long-ranged and thus non-trivial to treat in a simulation context. In this chapter we review our efforts to simulate such electrohydrodynamic systems using molecular dynamics particles to represent solvated ions, colloids and polymers. Our approach uses standard algorithms for the long-ranged electrostatic interactions and couples the particles to a Lattice-Boltzmann fluid in order to simulate the long-ranged hydrodynamic interactions. First we present this method in detail and follow with several example systems to which we have successfully applied the model. Then, we show that for a charged colloid in a monovalent salt solution the algorithm produces results which agree with both experimental data and theoretical calculations. Next, the algorithm is shown to reproduce the electrophoretic behavior of a single polyelectrolyte. The third system is a charged colloid grafted with oppositely charged polyelectrolytes, where we show that the system undergoes a mobility reversal as a function of the concentration of monovalent salt. The results are shown to be in good agreement with numerical calculations. The final system demonstrates how the Lattice-Boltzmann algorithm can be used to simulate the conductivity of a nanopore threaded with a DNA strand. The results are substantiated with the help of atomistic simulations.

1 Introduction

Typical aqueous solutions contain both negative and positive charges, often referred to as salt. When polymers, proteins, biological cells and colloids typically ionize when placed in an aqueous solutions. The acquired surface charge is most often negative. Modern micro- and nano-fluidic devices used to conduct experimental assays made out of silicon or glass also acquire a negative surface charge when placed in contact with an aqueous solution. From a simulation perspective, these prototypical soft matter systems pose a challenge as their conformations are largely governed by electrostatic interactions, which are long-ranged in nature and thus both non-trivial to implement and computationally expensive to calculate.

When these systems are subject to an external electric field the situation becomes even more complicated. Hydrodynamic interactions, which can be ignored in equilibrium, become of primary importance. In fact, as will be shown in this contribution, failing to properly take into account hydrodynamic interactions in such cases frequently results in not only erroneous quantitative values, but also in failing to capture the correct qualitative behaviour of the system. Despite this, many studies simply ignore hydrodynamic interactions due to the computational cost of fully taking them into account.

The systems where both electrostatic and hydrodynamic interactions can be roughly divided into three categories: 1) Electrophoresis, the movement of solvated objects due

to the application of an applied electric field, 2) Electroosmosis, the motion of fluid due to the application of an electric field tangential to a charged surface, 3) Conductance, the flux of electric charge through an aqueous solution. The inherent interplay of long-ranged electrostatic and hydrodynamic interactions in these phenomena make them a particular challenge to simulate. While a number of approaches exist for both types of interaction, the necessary simulation time often makes modeling such systems unfeasible in practice.

This contribution focuses on the application of Lattice-Boltzmann to the simulation of electrokinetic phenomena. Lattice-Boltzmann sacrifices atomistic detail in order to achieve a massive increase in speed. In fact, it has been shown that the relative speedup can be a factor 20 or higher using a CPU LB code[1,2]. A further speed increase of almost two orders of magnitude can be achieved by using a GPU LB[3]. The book chapter is structured as follows: First a brief introduction to the theory of hydrodynamics is presented in Section 2. Section 3 consists of a short review of the SEM. The Lattice-Boltzmann method is described in Section 4. Next we examine the electrophoretic mobility of a simple colloid in Section 5, followed by a linear polyelectrolyte in Section 6, and then a colloid grafted with polyelectrolytes in Section 7. Finally we take a detailed look at the conductance of a nanopore in Section 8. We end the contribution by providing a brief conclusion and discussion of future prospects in Section 9.

2 Hydrodynamics

This section provides an introduction to hydrodynamics, taken verbatim from[4]. In order to study the behaviour of particles within a solvent, a precise understanding of the fluid is needed. Due to the nonlinear and long-range nature of the forces involved, fluid mechanics is a highly complex topic and a complete description of hydrodynamic interactions easily fills books, such as Landau and Lifshitz[5]. The essential concepts and quantities that will be used within this chapter are briefly summarized in the following.

2.1 What are hydrodynamic interactions?

The Merriam-Webster's dictionary[a] defines *hydrodynamics* as "... a branch of physics that deals with the motion of fluids and the forces acting on solid bodies immersed in fluids and in motion relative to them".

Simply speaking, hydrodynamic interactions are long-range interactions arising within a fluid. An object that is moved through the fluid creates a movement of fluid particles around its surface called *flow*. This local disturbance of the fluid propagates through the medium. It can affect the movement of another object: if the flow created by the first object is strong enough, the second object might even be dragged along. The interaction necessary to mediate this force is called *hydrodynamic interaction*. Likewise, the presence of an object, *e.g.*, a wall, inside a moving fluid can influence the flow and divert it in such a way that other objects inside the fluid are not affected by it any more. The first object (the wall) *shields* the second object from the flow, thereby reducing the force exerted by the fluid on this particle. This effect is called *hydrodynamic shielding*.

In the following section, the theoretical framework that describes hydrodynamics from a physicist's point of view is presented.

[a] http://mw1.m-w.com/dictionary

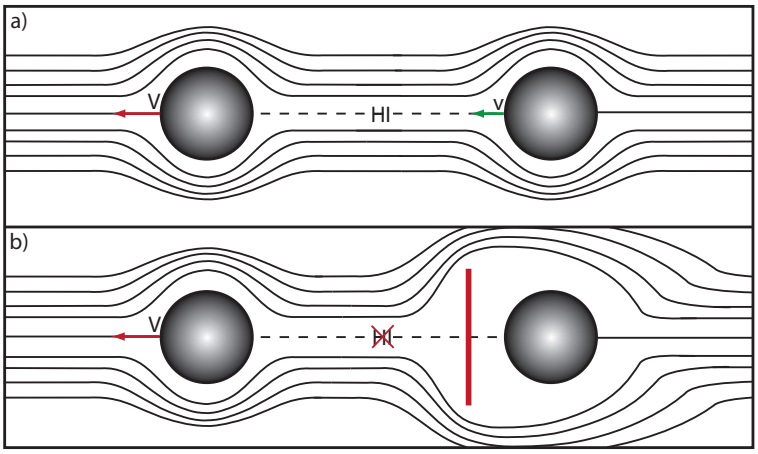


Figure 1: The motion of fluids and the forces acting on solid bodies immersed in them is called hydrodynamics. a) The flow created by a moving body can induce motion in a second body. b) An object placed in between the two bodies modifies and can even block the hydrodynamic interactions (HI). Figure reproduced from[4].

2.2 Navier-Stokes equation

In general, a fluid is composed of small particles or molecules that move around and collide with each other and with immersed objects following some conservation laws, namely the conservation of mass, momentum and energy known from classical mechanics. However, due to the large number of such particles within the fluid, a continuum assumption can be made. Only the combined effect of the individual fluid molecules is considered. In this case, the individual particle properties are replaced by averaged local quantities such as the density ρ and the flow velocity $\vec{v}$.

The dynamics of the continuous fluid are then described by the *Navier-Stokes equation*, which in its most general form can be written as:

$$\rho\left(\frac{\partial\vec{v}}{\partial t}+\vec{v}\cdot\nabla\vec{v}\right)=-\nabla p+\nabla\cdot\mathbb{T}+\vec{f}, \tag{1}$$

where $\vec{v}$ is the fluid velocity, ρ is the fluid density, p is the pressure, $\mathbb{T}$ is the general stress tensor, and $\vec{f}$ sums all forces acting on the fluid, be that gravity, centrifugal forces, etc.

As the Navier-Stokes equation only ensures conservation of momentum, an additional equation has to be enforced for conservation of mass:

$$\frac{\partial\rho}{\partial t}+\nabla\cdot(\rho\vec{v})=0. \tag{2}$$

Equation 1 can be simplified if additional assumptions about the nature of the fluid are taken into account. In the present study, only incompressible Newtonian fluids will be considered, for which the stress tensor $\mathbb{T}$ is proportional to the local strain $\nabla\vec{v}$ with the *dynamic viscosity* η as proportionality constant. In this case the Navier-Stokes equation

can be simplified:

$$\rho\Big(\underbrace{\frac{\partial \vec{v}}{\partial t}}_{\substack{\text{Unsteady}\\ \text{acceleration}}} + \underbrace{\vec{v}\cdot\nabla\vec{v}}_{\substack{\text{Convective}\\ \text{acceleration}}} \Big) = \underbrace{-\nabla p}_{\substack{\text{Pressure}\\ \text{gradient}}} + \underbrace{\eta\nabla^2\vec{v}}_{\text{Viscosity}} + \underbrace{\vec{f}}_{\substack{\text{Other}\\ \text{forces}}} . \tag{3}$$

2.2.1 Solving the Navier-Stokes equation

In general, the set of Navier-Stokes equations (Eq. 1) is composed by nonlinear partial differential equations in almost every real situation. Few exceptions exist, for example, in one dimensional flow, called also the Stokes flow (or creeping flow), for which Eq. 1 boils down to the single ordinary differential equation that can be easily solved analytically. Due to the nonlinearity, many problems are difficult or impossible to solve analytically. Furthermore, no closed-form solutions are proved to exist for the unsimplified Navier-Stokes equations.

However, many computational methods to solve the Navier-Stokes equations are known: finite volume methods (FVM)[b], finite element methods (FEM)[c], and finite difference methods (FDM)[d] to name a few. All of these methods discretize the whole fluid volume into smaller elements, and the partial differential equations are specialized to these elements making them numerically solvable. In general, the accuracy of the solution depends on the geometry of flow and the chosen discretization.

Mesoscopic fluid models will be introduced (see Section 4.7) that solve Navier-Stokes equations and allow for direct coupling to chosen simulation approach.

2.3 Important hydrodynamic quantities

When describing a specific fluid and its properties a whole zoo of hydrodynamic quantities can be used. So far, the fluid density ρ and the dynamic viscosity η have been introduced to describe the fluid itself. Often one is especially concerned with the ratio between the viscous force and the inertial force. This ratio is characterized by the *kinematic viscosity* ν:

$$\nu = \frac{\eta}{\rho}. \tag{4}$$

As a guideline, the density of water is roughly $\rho = 1\mathrm{g/cm}^3$, and its dynamic viscosity is $\eta \approx 1\mathrm{mPa{\cdot}s} = 1\mathrm{cP}$ (centi poise[e]). This leads to a kinematic viscosity of $\nu = 1\mathrm{mm}^2/\mathrm{s} = 1\mathrm{cSt}$ (centi stokes).

To characterise the fluid flow, several dimensionless physical quantities have been introduced. Below the four most commonly used ones are listed. All of them describe the ratio between different physical aspects of the fluid and can be used to distinguish different flow characteristics.

---

[b] `http://tinyurl.com/sdbp3`

[c] Ciarlet, Phillippe G. (1978). *The Finite Element Method for Elliptic Problems*. Amsterdam: North-Holland

[d] William F. Ames, *Numerical Methods for Partial Differential Equations*, Section 1.6. Academic Press, New York, 1977. ISBN 0-12-056760-1.

[e] named after Jean Louis Marie Poiseuille

The *Reynolds number* Re quantifies the ratio between the inertial forces ($V^2\rho D^2$) and the viscous forces (ηDV), where D is the relevant physical length scale, and it is a dimensionless quantity that is defined as follows:

$$\mathrm{Re} = \frac{\rho V D}{\eta}, \tag{5}$$

with V being the average fluid velocity. The Reynolds number is used to describe the onset of turbulent flow. Flows with a Reynolds number below a critical value ($\approx 10^4$ for pipes) are laminar, whereas flows with a higher Reynolds number might show turbulence.

For example, the Reynolds number of the blood flow in the aorta is Re $\approx 10^3$, the one of a person swimming is Re $\approx 10^6$, and finally the one of a large ship moving at full speed is Re $\approx 10^9$.

The second dimensionless number, the *Schmidt number* Sc, defines the ratio of momentum diffusivity ($\propto \nu$) and mass diffusivity D:

$$\mathrm{Sc} = \frac{\nu}{D} = \frac{\eta}{\rho D}. \tag{6}$$

As such, the Schmidt number relates the the convective matter transport to the diffusive transport.

In liquids, such as water, Sc ≈ 1000 and the momentum transport is much more prominent than actual matter transport. For example, while water waves collectively transport large momenta, the individual water molecule on average moves over very small distances.

Another dimensionless number is the *Knudsen number* Kn. It is a microscopic quantity defined as the ratio between the molecular mean free path length λ, the distance a molecule travels until the first collision, to a representative physical length scale L:

$$\mathrm{Kn} = \frac{\lambda}{L}. \tag{7}$$

The Knudsen number can be used to determine whether a fluid should be described by statistical mechanics or by continuum mechanics. If Kn > 1, *i.e.*, the mean free path of a molecule is comparable or larger to the relevant length scale, a continuum approximation is no longer valid.

All problems discussed in this chapter have Knudsen numbers smaller than one.

The *Mach number* Ma describes the speed of an object v moving through a fluid relative to the speed of sound c_s in it:

$$\mathrm{Ma} = \frac{v}{c_\mathrm{s}}. \tag{8}$$

Objects that move with a speed higher than Ma $= 1$ travel faster through a medium than the disturbance they cause to the medium. Obviously, this leads to strongly nonlinear effects that will not be treated in this work. All velocities in the systems of interest are subsonic with Ma < 1.

3 Electrokinetic Equations

In this section we provide an outline of the standard electrokinetic model, copied verbatim from[6]. When a charged particle immersed in a fluid is subject to a uniform applied

electric field $\vec{E}$, it moves in a process termed electrophoresis as illustrated in Fig. 2. The medium resists the particle's motion by exerting a drag force $\vec{F}_{\mathrm{d}}$. Usually, the particle is surrounded by counterions and in an electrolyte solution, an electric double layer (EDL) will form around it consisting of (i) the Stern layer consisting of the ions strongly attached to the surface, and (ii) the diffuse layer with a characteristic length scale called the Debye length λ_{D}. The counterions in the EDL also move under the influence of the electric field, dragging the fluid along with them. This generates an *electro-osmotic flow* (EOF) in the fluid, the velocity of which is the counterpart of the electrophoretic velocity of the particle. Since the counterions have a charge opposite to that of the particle, they move in the opposite direction and thus exert an extra hydrodynamic friction force on the particle, mediated by the surrounding solvent. This force is known as the retardation force $\vec{F}_{\mathrm{rt}}$.

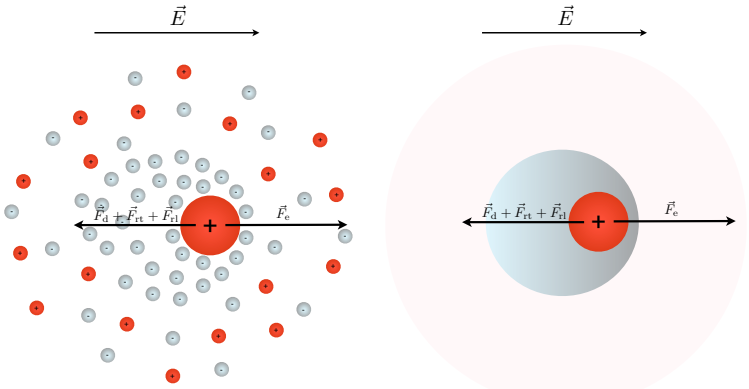


Figure 2: A sketch of the electrophoresis of a positively charged colloid (large red sphere). On the left-hand side the ions are depicted explicitly, the negative ones shown in gray and the positive ones shown in red, while the equivalent mean-field implicit ion approach is illustrated on the right-hand side. In the steady state, the electric force $\vec{F}_{\mathrm{e}}$ exerted on the colloid is canceled by the sum of the drag force $\vec{F}_{\mathrm{d}}$, the retardation force $\vec{F}_{\mathrm{rt}}$, and the relaxation force $\vec{F}_{\mathrm{rl}}$. Figure reproduced from[6].

The concurrent action of the applied field and that of the particle motion deforms the counterion cloud, creating a dipole which causes an electric retarding force to act on the particle called the relaxation force $\vec{F}_{\mathrm{rl}}$. The two retardation and relaxation forces are coupled; The asymmetry of the counterions cloud affects the retardation force, and the fluid's velocity distribution needs to be explicitly taken into account in the calculation of ion densities leading to relaxation force. In the steady state, the net force acting on the particle is zero and it acquires a drift speed v_{d}. At low field strengths, v_{d} is linearly proportional to the strength of the applied field $|\vec{E}| \equiv E$ and the proportionality constant is defined as the electrophoretic mobility:

$$\mu = \frac{v_{\mathrm{d}}}{E} \tag{9}$$

μ is the main quantity of interest in electrophoresis and also the one measured in experiments. It is related to ζ, the potential drop at the shear plane (introduced below) with respect to the bulk, and hence to the charge density of the object. When the fluid adjacent to a surface flows tangentially as a result of electrical forces, or shear mechanical forces,

or *etc.*, a thin layer, referred to as the Stern layer, is shown to adhere to the surface and become stagnant[7]. No hydrodynamic flow can develop within this layer and the fluid starts to shear beyond it. The imaginary surface separating this stagnant layer from the mobile part of the fluid is referred to as the *shear* or *slip* plane. Information about the value and distribution of the particle's charge, often unknown *a priori*, can be extracted by measuring μ and converting it to ζ. As will become clear in the course of this section, the conversion from μ to ζ, or vice versa, is not always straight forward and needs a theory which takes into account all the important properties of the system.

In order to calculate the electrophoretic mobility theoretically, a set of coupled non-linear differential equations, referred to as the electrokinetic (EK) equations, needs to be solved. They describe the motion of the fluid as well as that of the particle and the ions. In the following, the origin of the co-moving coordinate system is fixed on the particle. The force balance in the liquid can be described via the NS equations (Eq. 1), as introduced in Sec. 2. In the limit of low Reynolds numbers[f], the convective accelerations can be neglected and the NS reduces to Stokes equation which in the steady state ($\partial\vec{u}/\partial t = \vec{0}$) reads:

$$\eta\nabla^2\vec{u}(\vec{r}) - \nabla p(\vec{r}) - \sum_{j=1}^{N} n_j z_j e \nabla\psi(\vec{r}) = \vec{0}, \tag{10}$$

where n_j is the number density of the ion species j. The last term in this equation is the body force density due to the electrostatic interactions and introduces the retardation force in the equations. The electric potential ψ (including the applied field) satisfies Poisson's equation. At the low velocities typically encountered in electrophoresis, the fluid can in a good approximation be considered as incompressible, in which case the continuity equation simplifies to:

$$\nabla \cdot \vec{u}(\vec{r}) = 0. \tag{11}$$

The number of ions in the system is conserved:

$$\frac{\partial n_i}{\partial t} = \nabla \cdot (n_j \vec{v}_j) = 0, \tag{12}$$

where the ion velocities $\vec{v}_j$ can be found from the Nernst-Planck (NP) equation plus an advective term in the steady state:

$$\lambda_j(\vec{u}(\vec{r}) - \vec{v}_j) - z_j e \nabla\psi(\vec{r}) - k_\mathrm{B}T\nabla \log n_j = \vec{0}. \tag{13}$$

The first term of this equation is the convective contribution due to the fluid flow with λ_j being the drag coefficient of the ith ion species. The second term is the electric force due to the external field, and the last term is the diffusive contribution due to the ideal gas contribution to the chemical potential.

The boundary conditions needed to solve these equations are:

[f]The Reynolds number is a dimensionless quantity defined as the ratio of the inertial to the viscous forces and is used to determine the flow regime in a fluid or find similar flow patterns in different ones. Since it is the only parameter present in the non-dimensional form of the NS equation, the Reynold number characterizes the flow for a given geometry.

$$\begin{cases} \vec{u}(|\vec{r}| = r_\mathrm{s}) & = \vec{0}, \\ \vec{u}(|\vec{r}| \to \infty) & = -\mu\vec{E}, \end{cases} \tag{14}$$

associated with equations 10 and 12, where r_s is the position of the shear plane, and:

$$\begin{cases} \vec{v}_i(|\vec{r}| = r_\mathrm{s}) \cdot \hat{n} & = 0, \\ n_i(|\vec{r}| \to \infty) & = n_i^0(\infty), \end{cases} \tag{15}$$

for the ions, where $\hat{n}$ is the unit vector perpendicular to the surface and $n_i^0(\infty)$ is the equilibrium density of the ions in the bulk.

These equations govern electrokinetic phenomena and can, in principle, be used for solid surfaces of arbitrary shape. In this work, I will focus on the electrophoresis of spherical colloids. Solving this set of coupled nonlinear partial differential equations analytically is possible only for the most simple geometries or situations. However, for generic conditions the usual approach is numerical, and analytical solutions exist only for a few limiting cases which will be introduced below.

3.1 Hückel limit

In the absence of added salt, the only ions in the system are the intrinsic counterions produced by the dissociation of surface groups. For a charged sphere the electrical potential decreases as $1/r$ with distance and cannot compete with the increase in entropy which causes the counterions to "evaporate" from the surface. In an infinite volume or in the limit of no salt, the particle is merely a sphere with charge Q subject to an electric force $\vec{F}_\mathrm{e} = Q\vec{E}$. The only force opposing its motion is the Stokes viscous drag force exerted by the fluid $\vec{F}_\mathrm{d} = 6\pi\eta R\vec{v}$, where R is the radius of the particle and $\vec{v}$ its velocity. In the steady state, these two forces cancel each other and the particle moves with a draft velocity v_d:

$$\vec{v}_\mathrm{d} = \frac{Q}{6\pi\eta R}\vec{E}. \tag{16}$$

For low applied field strengths, the electrophoretic mobility in this salt-free limit, known as the "Hückel limit"[8,9], is then simply given by:

$$\mu = \frac{Q}{6\pi\eta R}. \tag{17}$$

Since the potential vanishes at infinity, the zeta potential in this case is merely the potential at the surface:

$$\zeta = \frac{Q}{4\pi\varepsilon R}, \tag{18}$$

such that the mobility in Eq. 17 can be rewritten as:

$$\mu = \frac{2\varepsilon}{3\eta}\zeta. \tag{19}$$

This analytical expressions derived for the Hückel limit can be used as an approximation for very low salt conditions where $\kappa R \ll 1$.

3.2 Helmholtz-Smoluchowski limit

In contrast to the Hückel limit, at very high salt concentrations the Debye layer is thin, $\kappa R \gg 1$. Most counterions accumulate in the close vicinity of the interface and "see" the particle as effectively flat. In other words, the local curvature at each point is much larger than the Debye length and the geometry is essentially planar. This is the well-known "Helmholtz-Smoluchowski limit". Under these conditions at low Reynolds numbers, the Stokes equation (Eq. 10), with a body force proportional to a tangentially applied electric field $\vec{E} = \{E_x, 0, 0\}$, yields the tangential fluid speed v_x as:

$$\frac{\mathrm{d}^2 v_x}{\mathrm{d}h^2} = -\frac{E_x \rho(h)}{\eta}, \tag{20}$$

where h is the vertical distance from the surface and we have assumed $\vec{\nabla} p = \vec{0}$ and a purely electric body force. Substituting in Poisson's equation we obtain:

$$\frac{\mathrm{d}^2 v_x}{\mathrm{d}h^2} = \frac{\varepsilon E_x}{\eta} \frac{\mathrm{d}^2 \phi}{\mathrm{d}h^2}. \tag{21}$$

This ordinary differential equation is straight forward to solve. Integrating both sides twice with respect to h and applying the boundary conditions $\mathrm{d}v_x/\mathrm{d}h = \mathrm{d}\phi/\mathrm{d}h = 0$ as $h \to \infty$, $v_x(h_\mathrm{s}) = 0$, $v_x(\infty) = v_\mathrm{eo}$, $\phi(h_\mathrm{s}) = \zeta$, and $\phi(\infty) = 0$, we find the Helmholtz-Smoluchowski (HS) equation for the EOF velocity:

$$v_x = v_\mathrm{eo} = -\frac{\varepsilon \zeta}{\eta} E_x. \tag{22}$$

The *electrophoretic* mobility in this limit is then:

$$\mu = -v_\mathrm{eo}/E_x = \varepsilon \zeta / \eta \tag{23}$$

. The HS equation describes the constant plug flow velocity profile outside the Debye layer. Inside the layer the charge density decreases exponentially with distance from the surface with a characteristic decay length equal to the Debye length $\lambda_\mathrm{D} = \kappa^{-1}$, and so does the velocity:

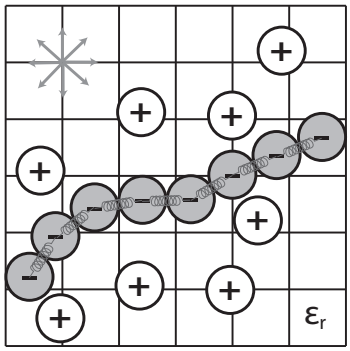
$$v_x(h) = v_\mathrm{eo} \left(1 - e^{-\kappa h}\right). \tag{24}$$

The HS formula applies to all particle shapes as long as the condition $\kappa R \gg 1$ is fulfilled and ζ is low, *i.e.* at very high salt concentrations the electrophoretic mobility is independent of the shape[10]. For high zeta potentials ($\zeta \gtrsim 2\, k_\mathrm{B}T/e \simeq 50\ \mathrm{mV}$) other effects such as the polarization of the Debye layer and surface conductance[7] become important and need to be taken into account. This is done either by solving the EK equations numerically or using approximate analytical expressions[7]. O'Brien provided such an expression which account for the diffuse-layer conductivity[11]:

$$\mu_\mathrm{red} \equiv \frac{3}{2} \frac{\eta e}{\varepsilon k_\mathrm{B} T} \mu = \frac{3}{2} \zeta_\mathrm{red} - \frac{6 \left[\frac{\zeta_\mathrm{red}}{2} - \frac{\ln 2}{z} \left(1 - \exp\left(-z \zeta_\mathrm{red}\right)\right) \right]}{2 + \frac{\kappa R}{1 + 3m/z^2} \exp\left(-\frac{z \zeta_\mathrm{red}}{2}\right)}, \tag{25}$$

where μ_red and $\zeta_\mathrm{red} = e\zeta/(k_\mathrm{B}T)$ are respectively the reduced electrophoretic mobility and the reduced $\zeta-$potential, and m is the dimensionless mobility of the ions ($m \simeq 0.15$ in aqueous solution). The O'Brien equation (Eq. 25) is a simplified form of an equation derived by Dukhin and Semenikhin[7] by neglecting terms of order $(\kappa R)^{-1}$.

4 The Hybrid Molecular Dynamics-Lattice-Boltzmann Model


+
+
+

4.2 Interaction potentials

All particles interact with one another via potentials that describe their physical nature.

The polyelectrolyte is comprised of N particles representing the individual monomers. A truncated, purely repulsive, Lennard-Jones or Weeks-Chandler-Anderson potential[14]

$$U_{\text{WCA}}(r_{ij}) = \begin{cases} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} + \frac{1}{4}\right] & \text{for } r_{ij} < r_{\text{cut}} \\ 0 & \text{for } r_{ij} \geq r_{\text{cut}} \end{cases} \tag{26}$$

is acting between all particles in the system determining their size. Here, r_{ij} is the distance between two interacting particles i and j, σ_{ij} sets the length scale, and ϵ_{ij} the energy scale of the interaction. The physical particle size is directly related to the parameter σ_{ij}, which is therefore also seen as the effective size of the particles. In this study, the cutoff $r_{\text{cut}} = \sqrt[6]{2}\sigma_{ij}$ ensures a purely soft repulsive interaction.

The N monomers are connected to a chain by finitely extensible nonlinear elastic (FENE) bonds[15]

$$U_{\text{FENE}}(r_{ij}) = \frac{1}{2}kR^2 \ln\left(1 - \left(\frac{r_{ij}}{R}\right)^2\right), \tag{27}$$

with stiffness k, maximum extension R, and r_{ij} again being the distance between the interacting monomers.

Together, these two potentials create a fully flexible self-avoiding polymer composed by N units. Stiff polymers with higher persistence length are modeled by including a bond angle potential

$$U_{\text{BA}}(\phi) = \frac{K}{2}(\phi - \phi_0)^2, \tag{28}$$

where the angle ϕ created by three connected particles is harmonically restored to the equilibrium angle ϕ_0 via an elastic constant K.

Unbound ions interact via a WCA potential, Eq. 26, that defines their physical size.

4.3 Continuum electrostatics

All particles in the system, *i.e.*, all N monomers and all ions, carry an electrical charge q_i that creates a non-bonded long-range interaction given by the Coulombic potential

$$U_{\text{C}}(r_{ij}) = \frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{\epsilon_{\text{r}} r_{ij}}. \tag{29}$$

Here, q_i and q_j are the effective charges on each particle, ϵ_0 is the permittivity of free space, and ϵ_{r} is the dielectric constant of the medium. When an explicit polar solvent is used, including ϵ_{r} is not necessary. However, for simulations with either implicit or non-polar solvent models, as the ones used in this study, ϵ_r is an effective dielectric constant that includes the screening effects due to the medium (*e.g.*, $\epsilon_{\text{r}} = 80$ for water).

When the system is simulated at a certain temperature T by using a thermostat, Equation 29 is commonly reformulated as

$$\frac{1}{k_{\text{B}}T}U_{\text{C}}(r_{ij}) = l_{\text{B}}\frac{\hat{q}_i\hat{q}_j}{r_{ij}}, \tag{30}$$

using the Bjerrum length l_{B} as defined by

$$l_{\mathrm{B}} = \frac{e_0^2}{4\pi\epsilon_0\epsilon_{\mathrm{r}}k_{\mathrm{B}}T}, \tag{31}$$

which represents the distance at which two unit charges experience an electrostatic potential that is equal to the thermal energy $k_{\mathrm{B}}T$.

4.4 Electrostatics in periodic boundary conditions

In this study, as well as in many others, *periodic boundary conditions* are used to approach bulk system behaviour within the limit of currently available computer systems. The main simulation cell is extended infinitely in all dimensions via copies of itself. In this way, it can be avoided that particles which are close to the boundary of the simulation cell, "feel" a different environment than particles in the center of the cell.

Unfortunately, the calculation of the electrostatic interaction is complicated by the periodic boundary conditions as not only the contribution of the real charges but also of the periodically repeated images has to be taken into account. Therefore, equation 29 has to be rewritten as follows:

$$U_{\mathrm{C,PBC}}(\vec{r}) = \frac{1}{4\pi\epsilon}\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{\vec{n}\in\mathbb{Z}^3}{}^{\dagger}\frac{q_i q_j}{|\vec{r}_{ij}+\vec{n}L|}, \tag{32}$$

where $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$, the difference between the position of the charges. The sum over $\vec{n}$ is taken over all simple cubic lattice points, $\vec{n} = (n_x L, n_y L, n_z L)$ with n_x, n_y, and n_z being integers. The † indicates that the $i = j$ term must by omitted for $\vec{n} = 0$ to avoid to take into account the interaction of a particle with itself.[g]

The use of Eq. 32 in order to compute electrostatic interactions is known as *direct sum method*, which, although simple to implement, suffers from a major drawback: the sum over $\vec{n}$ is an infinite series. This entails that when we want to evaluate the sum numerically we must perform a cutoff, *i.e.*, we assume that the contributions arising from larger $\vec{n}$ values can be neglected. Unlike the WCA interaction (Eq. 26), which is extremely short ranged, the Coulombic interaction is long ranged and only decays very slowly ($\propto \frac{1}{r}$). While a cutoff value is appropriate in one dimension[12], the long-ranged contributions are important in two or three dimensions and a too low cutoff value incurs large numerical errors. For this reason, inclusion of electrostatic effects can be computationally expensive and many sophisticated techniques have been developed to address this particular problem (see[16] and the references therein).

4.5 Ewald sum

The electrostatic force in a periodic system can be more efficiently calculated using an *Ewald sum*, which involves calculating the short range interactions in real space and the long range contributions in Fourier space[17,18].

[g] Equation 32 is only conditionally convergent in 3D. In other words, the value of the sum is not well defined unless one specifies the way we are going to sum up the terms (spheric, cubic, cylindric, etc.).

The algorithm is based on splitting the slowly convergent equation 32 into two series which can be computed much faster at a given accuracy by expressing $1/r$ as

$$\frac{1}{r} = \frac{f(r)}{r} - \frac{1-f(r)}{r}. \tag{33}$$

An usual choice is $f(r) = \text{erfc}(\alpha r)$, where α is called the *Ewald splitting parameter*, which result in the *Ewald formula* for the energy of the main cell

$$U = U^{(r)} + U^{(k)} + U^{(\text{self})} + U^{(\text{dipolar})} \tag{34}$$

where $U^{(r)}$ is called the real space contribution, $U^{(k)}$ is the reciprocal space contribution, $U^{(\text{self})}$ is the self-energy, and $U^{(\text{dipolar})}$ accounts for the dipolar correction.

In the case of performing a spherical sum over the lattice vectors $\vec{n}$, the expressions for these contributions are given by

$$U^{(r)} = \frac{1}{2}\sum_{i}^{N}\sum_{j}^{N}\sum_{\vec{n}\in\mathbb{Z}^3}{}^{\dagger} q_i q_j \frac{\text{erfc}(\alpha|\vec{r}_{ij}+\vec{n}L|)}{|\vec{r}_{ij}+\vec{n}L|} \tag{35}$$

$$U^{(k)} = \frac{1}{2L^3}\sum_{\vec{k}\in\mathbb{K}^3,\vec{k}\neq 0}\frac{4\pi}{k^2}\exp\left(-k^2/4\alpha^2\right)\sum_{i}^{N}\sum_{j}^{N} q_i q_j \exp(-i\vec{k}\cdot\vec{r}_{ij}) \tag{36}$$

$$U^{(\text{self})} = -\frac{\alpha}{\sqrt{\pi}}\sum_{i}^{N} q_i^2 \tag{37}$$

$$U^{(\text{dipolar})} = \frac{2\pi}{(1+2\epsilon)L^3}\left(\sum_{i}^{N} q_i r_i\right)^2 \tag{38}$$

where $\mathbb{K}^3 = \{2\pi\vec{n}/L : \vec{n}\in\mathbb{Z}^3\}$.

In practice the sums for $U^{(r)}$ and $U^{(k)}$ are evaluated performing cutoffs given by r_{cut} and k_{cut}. Typical implementations, as the one we are going to use, assume the *minimum image convention*, *i.e.*, , $r_{cut} < L/2$ and therefore $\vec{n} = 0$ in the expression for $U^{(r)}$.

The advantage of the Ewald sum is two-fold. On the one hand, the convergence of Equations 35 and 36 is faster than the Equation 32 making it possible to define a more efficient cutoff for a given accuracy. On the other hand, it is possible to derive exact estimates for the error incurred by applying a certain cutoff.

4.6 Fast calculation of electrostatics

Although the Ewald sum method represents a substantial improvement with respect to the direct sum, in many cases it is not as fast as one would desire. This drawback is specially notorious when we are dealing with systems containing a large number of charged particles and is mainly due to the fact that the computer time in Ewald sums scales with the number of particles as $\mathcal{O}(N^2)$, or in the best versions as $\mathcal{O}(N^{3/2})$. These scalings entail that the computer time grows very quickly with the number of particles.

The approach is efficiently sped up by the use of fast Fourier transformation (FFT) in commonly used algorithms: the so-called particle-particle-particle mesh (P3M)[19,20], particle mesh Ewald (PME)[21], and smooth PME (SPME)[22] methods. See[23] for a review of the individual methods and a detailed comparison. Further interesting approaches to

solve this problem are the (fast) multipole methods[24] and the local reaction field (LRF) method[25]. A number of fast electrostatic algorithms have been recently implemented into the ScaFaCoS library[26].

In this study, the P3M algorithm is used that scales like $\mathcal{O}(N \log(N)) \sim \mathcal{O}(N)$. Deserno *et al.*[27] derived error estimates that predict the achieved accuracy of the method based on the parameters chosen. This and other error estimates[28,29] are used in an automated tuning routine that checks all parameter combinations that yield the desired accuracy α for the given system and selects the set that uses the least computation time. Recently we also showed how to convert an SPME algorithm into the preferable P3M[29]. As a last note we would like to note that all so far mentioned Coulomb solvers assume a uniform dielectric background in which the charges reside. This is mostly true for bulk water, but can be dangerous for charges near interfaces or inhomogeneities. For this, also specialized Coulomb routines exist[30–35].

4.7 Mesoscopic fluid models

An alternative approach to include hydrodynamic interactions is to model an *explicit fluid* where fluid particles are interacting via a WCA potential (Eq. 26). This conceptually obvious approach is easy to implement and preserves long-range hydrodynamic interactions. In practise, however, as the trajectory of every fluid particle has to be calculated explicitly, much computing time is spent on details of the fluid which are often not of interest. For this reason, several modeling techniques to maintain the hydrodynamic interactions within the fluid while neglecting the costly computational details of the fluid motion have been developed. The description of the hydrodynamic effects in these approaches lies in between the use of an explicit description of fluid molecules and the solution to the continuous Navier-Stokes equations (Eq. 1). In particular, a clever use of conservation laws allows these methods to employ local algorithms to recover the solution to the hydrodynamic equations in the large-scale/long-time limit, bypassing at the same time the molecular detail of the fluid and thus reducing dramatically the computational cost of the simulation.

Nowadays, the most commonly used methods are Multiparticle Collision Dynamics (MPC) – also called Stochastic Rotation Dynamics (SRD) – [36–38], Dissipative Particle Dynamics (DPD)[39–41], and Lattice-Boltzmann (LB)[42–45]. There are several recent publications reviewing the different mesoscopic fluid models:[46–48].

4.7.1 Comparison of mesoscopic fluid models

All of the above methods use a simple (but sufficient) model to describe fluid dynamics, and can also act as thermostats that define the local temperature when coupled to MD particles. As such, they are well suited to study the polyelectrolyte electrophoresis as they can dissipate the energy transferred to the system by the applied electric field.

Due to their conceptual differences, these methods use different types of parameters to describe the fluid, which results in a different suitability for specific problems. For example, while the fluid viscosity is directly accessible in LB methods, it becomes a combination of different parameters and can only be controlled indirectly in DPD[49]. Similarly, while DPD and LB can only approximate the continuous-time dynamics of the fluid when the discrete time step is small, SRD is proven to yield correct long-time hydrodynamic

interactions for any step size. However, SRD's transport properties depend explicitly on the chosen time step[50]. More differences between the methods arise if confined fluids or interactions with large obstacles or particles are studied. Here, the ability to treat different boundary conditions becomes important, which is covered in detail in the literature[51,52].

All three methods share a similar computational efficiency, and computation times depend mainly on the implementation, on the computer system, and also on the investigated system, see[53] for an example. However, the speedup over explicit fluid simulations can be a factor 20 or higher[1,2].

This study uses the Lattice-Boltzmann method that is briefly reviewed in the following section. Related studies with DPD[54,55], MPCD[56,57], or a LB-PB lattice solver[58–60] to cover electrokinetic phenomena exist.

4.8 The Lattice-Boltzmann method

The Lattice-Boltzmann (LB) method is based on a solution of the discretized Boltzmann transport equation, which by means of a Chapman-Enskog expansion leads to the Navier-Stokes equation (Eq. 1) in the incompressible limit[43,45].

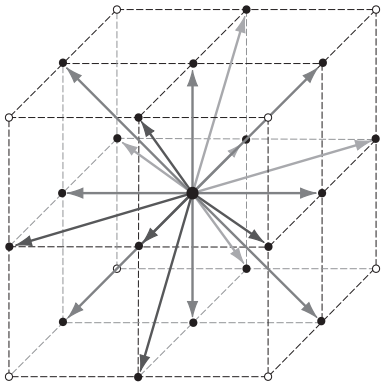


Figure 4: The 18 velocity vectors $\vec{c}_i$ for a D3Q18 lattice. From the central grid point, the velocity vectors point towards all 18 nearest neighbors marked by filled circles. Figure reproduced from[4].

The main quantities in the LB approach are the particle densities $n_i(\vec{r},t)$ at each point $\vec{r}$ on a spatially discrete grid at time t, rather than the fluid particles in the volume. It is an inherently statistical approach, where discrete velocities $\vec{c}_i(a/\tau)(i=1,\ldots,b)$ are associated to the particle densities $n_i(\vec{r},t)$, with a being the grid spacing, τ the time step of the discretization, and the $\vec{c}_i$ vectors pointing towards the ith of b next neighbors on the grid. Each of the $n_i(\vec{r},t)$ can be interpreted as the fraction of fluid that will move with the i-th discretized velocity at time t and position $\vec{r}$.

The particle densities $n_i(\vec{r},t)$ define the hydrodynamic quantities of interest, the mass density $\rho(\vec{r},t)$, the fluid momentum $\vec{j}(\vec{r},t)$, and the fluid velocity $\vec{u}(\vec{r},t)$, as given by the

following equations:

$$\rho(\vec{r},t) = \sum_i n_i(\vec{r},t) \tag{39}$$

$$\vec{j}(\vec{r},t) = \sum_i n_i(\vec{r},t)\vec{c}_i \tag{40}$$

$$\vec{u}(\vec{r},t) = \vec{j}(\vec{r},t)/\rho(\vec{r},t). \tag{41}$$

The LB method consists of two steps, a *collision step* and a *streaming step*. During the collision step the particle densities at each lattice site are updated. In the streaming step, the updated particle densities are propagated over the grid according to the associated velocity vectors.

The evolution of $n_i(\vec{r},t)$ in the collision step is described by the Lattice-Boltzmann equation:

$$n_i(\vec{r}+\vec{c}_i a, t+\tau) = n_i(\vec{r},t) + \sum_{j=1}^{b} L_{ij}(n_j(\vec{r},t) - n_i^{eq}(\rho,\vec{u})), \tag{42}$$

where the last term expresses the relaxation towards a local pseudo-equilibrium. The discretized Boltzmann equation provides a generic description for the time evolution of the particle densities, but there is freedom in the choice for the actual form of the collision integral L_{ij}. A common formulation is the Bhatnagar-Gross-Krook (BGK) approximation[61] with a diagonal $L_{ii} = \tau/\tau_\mathrm{r}$, where τ_r is a phenomenological relaxation time. It prescribes the timescale for the relaxation towards the equilibrium density $n_i^{eq}(\rho,\vec{u})$, which can be written as

$$n_i^{eq}(\rho,\vec{u}) = \rho a^{c_i}\left(1 + \frac{\vec{u}\cdot\vec{c}_i}{c_\mathrm{s}^2} + \frac{(\vec{u}\cdot\vec{c}_i)^2}{2c_\mathrm{s}^4} - \frac{u^2}{2c_\mathrm{s}^2}\right) \tag{43}$$

with c_s being the speed of sound.

However, in ESPResSo, following[44,1,62], a general form for L_{ij} was used, where the matrix structure is derived from physical and numerical arguments. The method is generally referred to as the Multi-Relaxation Time (MRT) version of the Lattice-Boltzmann developed by d'Humières[63]. In this formulation the collision occurs in mode space. In mode space the first mode is the mass, followed by three momentum mode as given in equation 39. The next three modes represent the bulk stress of the fluid, and are followed by three modes representing the shear stress of the fluid. The bulk modes are linearly relaxed with a relaxation time $\tau_\mathrm{b} = \tau/\lambda_\mathrm{b}$ while the shear modes are relaxed with a relaxation time $\tau_\mathrm{s} = \tau/\lambda_\mathrm{s}$. The 9 higher order modes, which have no obvious physical meaning, are called ghost modes and are not relaxed as it is unclear how to properly relax them. For a more detailed explanation we recommend the doctoral thesis of Ulf D. Schiller[62].

One can show via Chapman-Enskog expansion[h] that this model leads to the Navier-Stokes equations in the limit of small Knudsen and Mach numbers. From this, the relation

[h]The Chapman-Enskog expansion is essentially an asymptotic analysis in time and space with $t_1 = \epsilon t$ and $r_1 = \epsilon r$ followed by a Taylor expansion of n_i to second order in ϵ. Using that the conservation laws for mass and momentum must hold at every order in ϵ together with the substitution of the expansion of n_i into the Boltzmann equation 42 leads to the hydrodynamic equations when collecting terms in orders of ϵ.

between the non-trivial eigenvalue λ of L_{ij} and the kinematic viscosity ν follows:

$$\eta_{\mathrm{s}} = -\frac{\rho a^2}{3\tau}\left(\frac{1}{\lambda_{\mathrm{s}}} + \frac{1}{2}\right) \tag{44}$$

$$\eta_{\mathrm{b}} = -\frac{2\rho a^2}{3\tau}\left(\frac{1}{\lambda_{\mathrm{b}}} + \frac{1}{2}\right) \tag{45}$$

which can be used to specify the shear, η_{s}, and bulk, η_{b}, viscosity as an explicit input parameter for the LB method.

In order to thermalize the fluid under electrophoretic conditions, a stochastic term

$$n_i'(\vec{r},t) = -D_q \sum_{\alpha\beta} \sigma'_{\alpha\beta} c_{i\alpha} c_{i\beta} \tag{46}$$

fulfilling fluctuation-dissipation relation is added to the Lattice-Boltzmann equation (42)[44]. It has recently been shown that this noise has to be applied to all non-conserved modes of the collision operator to prevent poor thermalization on smaller length scales, and not only to the elements of the viscous stress tensor[64,65].

4.8.1 Coupling LB to coarse-grained molecular dynamics

On the length and time scales important in polyelectrolyte electrophoresis, the physical observables do not depend on the microscopic details of the coupling between the LB fluid and the coarse-grained molecular dynamics particles, as long as it is assured that the hydrodynamic interactions within the fluid evolve on time scales faster than the diffusion time scale of the monomers. Therefore, it is not necessary to resolve the shape of the particles for the fluid, but they can instead be treated as point particles. This approach was originally suggested by Ahlrichs *et al.*[1].

In analogy to the Stokes friction for a sphere in viscous fluid, the force on a particle exerted by the fluid is assumed to be proportional to the difference between the monomer velocity $\vec{v}$ and the fluid velocity $\vec{u}$ at the monomer's position

$$\vec{F}_{\mathrm{D}} = -\Gamma_{\text{bare}}\left(\vec{v}(t) - \vec{u}(\vec{r},t)\right). \tag{47}$$

Here, the friction coefficient Γ_{bare} determines the strength of the interaction between fluid and particles. In order to conserve total momentum of fluid and monomers, the opposite force has to be assigned to the fluid in the particles' cell fulfilling

$$-\vec{F}_{\mathrm{D}}/a^3 = \frac{\Delta\vec{j}}{\Delta t} = \sum_i \Delta n_i(\vec{r},t)\vec{c}_i \frac{\rho}{a^2\tau\Delta t}. \tag{48}$$

As the fluid velocity $\vec{u}$ is only defined at the grid sites, $\vec{u}(\vec{r},t)$ has to be interpolated to the particle position. This is done by a simple linear interpolation with the neighboring grid nodes, as indicated in Figure 5.

Additionally to the dissipative coupling, a stochastic random force is added to model the thermalizing properties of the fluid:

$$\vec{F}_{\mathrm{R}} = \vec{\zeta}_i(t). \tag{49}$$

As for simulations using a Langevin thermostat, $\vec{\zeta}_i(t)$ has to fulfill the following equations:

$$\langle\vec{\zeta}_i(t)\rangle = 0 \quad \text{and} \quad \langle\vec{\zeta}_i(t)\cdot\vec{\zeta}_j(t')\rangle = 6\Gamma_{\text{bare}}k_{\mathrm{B}}T\delta_{ij}\delta(t-t'). \tag{50}$$

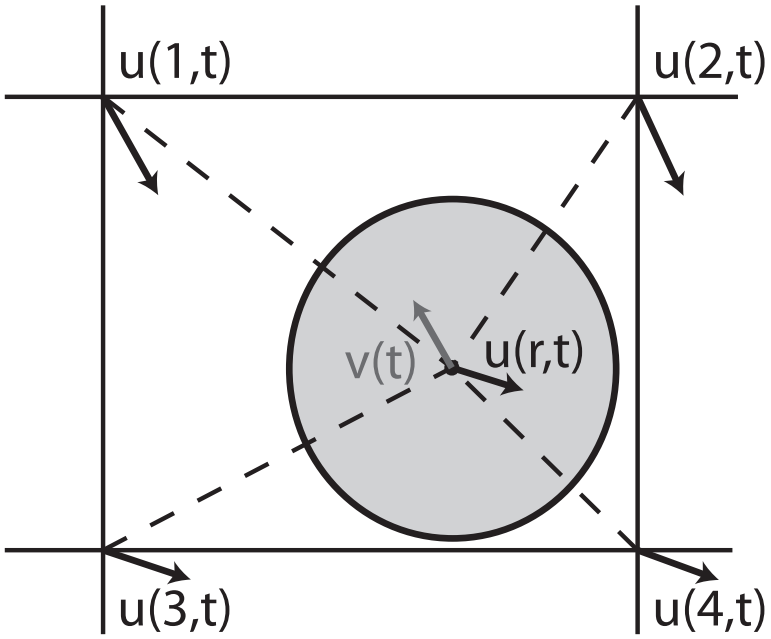


Figure 5: The coupling scheme between fluid and particles is based on the interpolation of the fluid velocity $\vec{u}$ from the grid nodes marking the cell the particle is in to the actual position $\vec{r}$ of the particle. This is done by linear interpolation. The difference between the actual particle velocity $\vec{v}(t)$ and the interpolated velocity $\vec{u}(\vec{r},t)$ is used in the momentum exchange of Equations 47 and 48. Figure reproduced from[4].

Again, the added momentum has to be transferred to the fluid following Eq. 48.

Both $\vec{F}_{\mathrm{D}}$ and $\vec{F}_{\mathrm{R}}$ are added to the total force $\vec{F}_{\mathrm{T}}$ before integrating the equations of motion for the particles. At the same time, the fluid is updated according to the LB method. The momentum exchange between particles and fluid correctly reproduces long-range hydrodynamic interactions at relatively low computation costs compared to an explicit fluid.

4.8.2 Validity of the LB Coupling Scheme

In this subsection (taken verbatim from[66]), we take a deeper look at the coupling scheme introduced in the previous subsection. We note that the considered coupling mechanism is not over-damped, but resolves the inertial time scale. It is helpful to quantify this scale for an illustrative case. Assuming a colloid of $R = 1$ μm radius with the density of water immersed in water, the characteristic time scale of the decay of the velocity to the velocity is

$$m/\Gamma_{\mathrm{bare}} = \rho\frac{4\pi}{3}R^3/6\pi\eta R \approx 200\,\mathrm{ns}. \tag{51}$$

In comparison, the time to diffuse over its own radius is

$$r^2/Dk_{\mathrm{B}}TR^2/6\pi\eta R \approx 5\,\mathrm{s}. \tag{52}$$

This time scale is a factor of approximately 10^7 larger than the inertial time scale. Accordingly, the Schmidt number is 10^7. Matching both numbers independently is thus very difficult, as the number of time steps must be at least equal to the ratio of the two. Using the Lattice-Boltzmann algorithm for colloidal particles is only be practical following the idea of *telescoping of time scales*[67]. If two time scales in a physical system are sufficiently separated, a simulation method can be applied where the time scales are closer together but still sufficiently separated. Practically, a much smaller Schmidt number may be as good as the correct one, which allows a significant reduction in the number of time steps. The clear advantage of using an algorithm where the inertial time scale is resolved, is that hydrodynamic interactions can be achieved by purely local operations.

Interestingly the friction coefficient Γ_{bare} in Eq. 47 is not identical to the inverse mobility of a particle. When a force $\vec{F}_{\text{ext}}$ is applied to a particle also the underlying fluid starts to move, as it is accelerated by the friction force. A stationary state is reached, when the applied forces and the friction force cancel, and then the applied force is completely transferred to the fluid. The particle velocity will then be given by

$$\vec{v} = \vec{F}_{\text{ext}}/\Gamma_{\text{bare}} + \vec{u}\left(\vec{r}\right). \tag{53}$$

Note, that due to the finite resolution of the grid, the fluid velocity at the particle position remains finite, in contrast to the divergent Stokeslet in the continuum case. Hence, if the resolution of the LB fluid is raised, the velocity at the position of the particle can be expected to be proportional to the resolution, as the flow field, decaying like $1/r$, is finer resolved. Furthermore in the Stokes equation the fluid velocity is proportional to the inverse of the viscosity. These dimensional arguments lead to a particle mobility of the form

$$\vec{\mu} = \frac{1}{\Gamma_{\text{bare}}} + \frac{g}{\eta a}, \tag{54}$$

which was suggested by Ahlrichs and Dünweg[1]. g is a numerical constant, which is supposedly independent of all other parameters. We performed the same check of this formula as Ahlrichs and Dünweg, and could confirm this behaviour for sufficiently high viscosities. As the LB method only fulfills the Stokes equation in the limit of large lengths, the flow field on short distances may deviate from the $1/\eta a$ scaling discussed above. This is why the numerical value of g also depends weakly on the viscosity. For the D3Q19 method with the first order coupling, g is around 0.04.

We define the hydrodynamic radius R_{H} of a particle coupled by this method as the radius of a sphere with the same mobility in the same medium. This leads to the following expression

$$\frac{1}{R_{\mathrm{H}}} = \frac{6\pi\eta}{\Gamma_{\text{bare}}} + \frac{6\pi g}{a}. \tag{55}$$

Thus for all choices of η and Γ_{bare}, the hydrodynamic radius can be at maximum

$$R_{\mathrm{H}} \leq \frac{a}{6\pi g} \approx 0.75a. \tag{56}$$

This point is very important from a practical perspective. Making the lattice of the LB finer does not leave the physics unchanged. On the contrary, it weakens the hydrodynamic interactions, as the hydrodynamic radius is decreased.

As a further test, we investigated the conductance of an electrolyte. Onsager[68], based on work of Debye and Hückel, derived a model for how the conductance of an electrolyte depends on the concentration. His theory applies for sufficiently small concentrations, and contains two effects, *relaxation* and *retardation*. First, the radial distribution of ions is skewed by the applied electric field, the relaxation effect. This leads to a reduction of the field strength at the position of the ion, and hence a reduction of the conductance. Second, ions of opposite charge moving in opposite directions are subject to hydrodynamic interactions which also slow them down, the retardation effect. In Onsager's calculation, hydrodynamic interactions are taken into account on the Oseen level. The results of this calculation is known as the Debye-Hückel-Onsager limiting law and is contained *e.g.*, in many physical chemistry books. For a 1:1 electrolyte, it can be brought to a particularly simple form, where it reads as

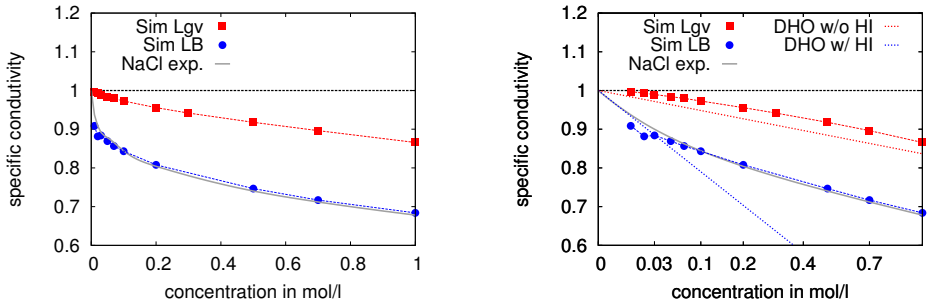
$$\sigma = 2c\mu_0 \left(1 - 0.07\frac{\lambda_\mathrm{B}}{\lambda_\mathrm{D}} - \frac{R_\mathrm{H}}{\lambda_\mathrm{D}}\right), \tag{57}$$

where c is the salt concentration. It is apparent, that the retardation term vanishes for vanishing hydrodynamic radii.

For small c, the Debye-Hückel-Onsager law is exact. Interestingly, this means that the conductivity can not be expanded in a Taylor series in c, as the slope at $c = 0$ is infinite, but must be expanded in terms of $\sqrt{c}$. Different empirical extensions have been proposed to extend the range of applicability of this formula to higher salt concentrations.

We performed simulations of a monovalent electrolyte at room temperature represented by spherical particles interacting with the WCA potential with $\sigma = 0.45$ nm and $\epsilon = k_\mathrm{B}T$. The hydrodynamic radius were set to be commensurate with that of K and Cl ions to 0.15 nm. This was done by using a lattice constant of 0.34 nm. This length was chosen as the unit of length in the simulations. The mass of both cations and anions was chosen as unity. Using the mean experimental mass this fixes the unit of time of our simulations to 1.5 ps. We chose a viscosity of 0.8 in simulation units, corresponding to 0.11 mPa s. This is a factor of 10 smaller than real water, but the resulting diffusion constant is 1.3 $10^{-8}\mathrm{m}^/\mathrm{s}$ also a factor of 10 larger. The Schmidt number is even smaller by a factor of 100, as we have reduced the viscosity and enhanced the diffusion constant. The box shape is cubic with an edge length of 20 simulation units, or 6.68 nm. The choice of the parameters is convenient from a practical perspective, as the diffusion constant is relatively high, and therefore simulated systems exhibit a high signal-to-noise ratio. On the other hand, the fluctuation-dissipation theorem is fulfilled sufficiently precisely to derive unique transport coefficients. It is, however, not unique, and from our perspective further investigations of the role of the parameters are necessary.

In Fig. 6 the result is depicted and compared to experimental results to NaCl from[69]. Furthermore we compare the results to simulations without hydrodynamic interactions using Langevin dynamics. The friction coefficient in the Langevin dynamics simulations are chosen such that the mobility of an isolated particle is matched. The conductance of the experimental data is normalized by its value at infinite dilution, and the simulation data is normalized by the mobility of isolated particles.



(a) (b)

Figure 6: Specific conductance of an electrolyte simulated with Langevin dynamics and the LB method. (a) Plotted with a linear concentration axis, (b) axis scaled by the square root of the concentration. Simulation data with Langevin dynamics are shown as squares, and data obtained with the LB method are shown as circles. Experimental data, interpolated with splines, is shown as grey lines. The result of the Debye-Hückel-Onsager (DHO) formula is added in the right plot as dashed lines, both with (blue) and without the hydrodynamic contribution (red). Figure reproduced from[66].

The agreement with the experimental data is very good. The simulation data captures the experimental curve almost to the statistical accuracy of the simulation data. The relaxation effect is weaker than the retardation effect, but constitutes an important contribution. For small electrolyte concentrations we find considerable deviations, that are caused by the finite size of our simulation box. For the smallest considered salt concentration only four ion pairs were used. Such a good agreement between simulation and experiment is somewhat surprising, and quite encouraging. Our simulation data not only captures low concentration as the Debye-Hückel-Onsager law, but can describe the experimental data over the entire concentration range.

4.8.3 Technical remarks on LB

The implementation used in this study allows to specify the fluid and the coupling via five parameters:

- the average fluid mass density ρ,
- the kinematic viscosity ν,
- the grid spacing a,
- the time step τ used for the fluid evolution,
- and the coupling parameter Γ_{bare}.

Whereas the first two parameters define the physical properties of the fluid, and the following two parameters detail the numerical discretization, the coupling parameter has no predefined meaning. However, it can be related to the single particle friction coefficient Γ_{bare} of the implicit fluid model via equation 55

		Description	Typical values used
General	N	length of polymers in monomers	$1 \ldots 256$
	n_c	number of chains	$1 \ldots 20$
	c_S	concentration of additional salt	$0 \ldots 1\mathrm{M}$
	σ_0	length scale of the system	$2.5 \ldots 4.2\,\text{Å}$
	$k_\mathrm{B}T$	energy scale of the system	1.0
	m_0	mass scale of the system	1.0
Interaction potentials (Section 4.2)	σ_{ij}	length scale of the WCA potential	1.0
	ϵ_{ij}	energy scale of the WCA potential	0.25
	k	spring constant for the FENE springs connecting the PE monomers	30.0
	R	maximal extension of the FENE springs	1.5
	K	harmonic constant of the bond potential to increase the persistence length	$0 \ldots 30.0$
	ϕ_0	equilibrium angle of the bond potential	$0°$
	Δt	time step for the integration of motion	0.01
Electrostatics (Section 4.3)	q_i	charges of the particles	1.0
	l_B	Bjerrum length ($\approx 7.1\,\text{Å}/\sigma_0$)	$1.7 \ldots 3.4$
	α	accuracy requirement for electrostatics	$1e^{-4}$
Hydro-dynamics (Section 2)	ρ	density of the mesoscopic LB fluid	$0.5 \ldots 3$
	ν	kinematic viscosity	$0.3 \ldots 3.0$
	a	grid spacing of the lattice	1.0
	Γ_bare	coupling constant for the particles	$20.0 \ldots 50$
	τ	time step for the update of the fluid	$0.01 \ldots 0.05$

Table 1: Summary of the parameters typically employed in hybrid LB-MD simulations, including the range of values used in this study. Parameters without physical units are specified in reduced MD units.

Finally, it remains to be pointed out that the parameters specifying the length and time scales of the molecular dynamics simulation, σ_{ij} and Δt, and of the LB fluid, a and τ, can be chosen independently of one another. However, as the grid spacing a defines the length scale on which hydrodynamic interactions are correctly resolved, it should be similar or smaller than σ_{ij}. On the other hand, one can confirm easily that τ can be chosen up to 5-times larger than Δt without a significant loss in accuracy for the systems investigated in this work. As the updating of fluid is the most time-consuming step during simulation, this provides a significant speed-up in computation time.

4.9 Typical system parameters

The typical parameters given in Table 1 can be specialized to match the physical properties of a target system.

Unless indicated, all simulation parameters are given in reduced units in the energy scale $k_\mathrm{B}T$, the mass scale m_0 and the relevant length scale σ_0 that is used to match the model to a specific physical system.

4.10 Determining transport coefficients in simulations

The simulation model is used to determine two different transport coefficients for the model polyelectrolyte that are likewise determined in the associated experiments.

4.11 Diffusion

The diffusion coefficient D characterises the thermal motion of the polyelectrolyte. It is obtainable from the simulation trajectory of the polyelectrolyte chain, by measuring the slope of the center of mass' mean-square displacement

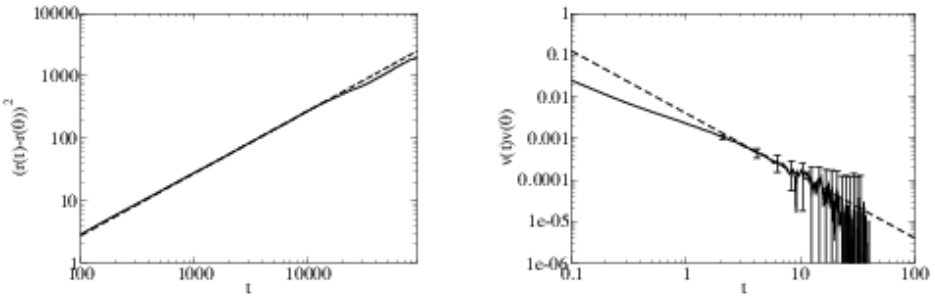
$$D = \frac{\left\langle [\vec{r}_{\text{cm}}(t) - \vec{r}_{\text{cm}}(0)]^2 \right\rangle}{6t},$$

where $\vec{r}_{\text{cm}}$ is the position of the center of mass, and t is the time. The angular brackets $\langle \ldots \rangle$ indicate the averaging over many configurations.

Alternatively, the diffusion coefficient D can be obtained from the integration of the velocity auto-correlation function of the center of mass

$$D = \frac{1}{3} \int_0^\infty \langle \vec{v}_{\text{cm}}(t) \cdot \vec{v}_{\text{cm}}(0) \rangle \, dt. \tag{58}$$

Here, $\vec{v}_{\text{cm}}$ is the center of mass velocity of the polyelectrolyte at a given time. Again, the angular brackets $\langle \ldots \rangle$ indicate the averaging over many configurations.



(a) (b)

Figure 7: The diffusion coefficient D of a polyelectrolyte chain of length $N = 32$ is determined (a) via a fit (dashed line) to the linear part of the mean square displacement, which yields a diffusion coefficient of $D = 0.0045 \pm 0.0002$ via Equation 4.11, (b) and via fitting a $t^{-3/2}$ power law to the long-time tail of the center of mass auto-correlation function and using Equation 58, which results in $D = 0.0041 \pm 0.0005$. Figure reproduced from[4].

The accuracy of both methods depends on the number of statistically independent data samples. In Figure 7, sample graphs to determine the diffusion of a polyelectrolyte chain of length $N = 32$ using Equations 4.11 and 58 are presented. Since simulations with hydrodynamic interactions are computationally very demanding, the achievable accuracy

is limited. The errors are determined from the statistical fluctuations of the data and the uncertainty in the fit parameter.

In Figure 7a, the diffusion is obtained from a fit to the linear part of the mean square displacement. This yields a diffusion coefficient of $D = 0.0045 \pm 0.0002$ in simulation units. To calculate the integral in Equation 58, a fit to the slowly decaying long-time tail of the center of mass' velocity auto-correlation function has to be obtained as shown in Figure 7b. Here, the theoretical predicted functions are used to match the long-time tail: without hydrodynamic interactions, an exponential decay of the velocity correlations is expected, whereas with hydrodynamic interactions the correlation function decays with $t^{-3/2}$. The figure shows the more interesting case with hydrodynamic interactions. The combined results of simulation data and long-time fit are integrated and a diffusion coefficient of $D = 0.0041 \pm 0.0005$ is obtained, which is in agreement with the corresponding value obtained from the mean square displacement.

Both methods are strictly equivalent for classical systems, but for the remainder of this work, the integral method is used to obtain the diffusion coefficient as one can use a similar formulation to obtain the mobility of the polyelectrolyte (see Equation 60), and thus can determine both quantities without additional computational effort.

4.12 Electrophoretic mobility

The second transport coefficient of interest is the electrophoretic mobility μ. It characterises the motion of the polyelectrolyte in an external electric field.

In capillary electrophoresis experiments[70], the electrophoretic mobility of the solute is determined by the ratio of the velocity to the applied electric field. This method can be directly transferred and applied to computer simulations. The external electric field E is modeled by a constant force proportional to the electric charge acting on the particles in the solution. This causes a directed motion with a certain velocity v. From this one can obtain the electrophoretic mobility

$$\mu = \frac{v}{E}. \tag{59}$$

Transforming this in a straightforward way to simulations is not always advantageous. Polyelectrolyte mobilities in free solution are of the order of $4 \cdot 10^{-8}$ m^2/Vs[71–74]; under experimentally used electric fields of up to 1000 V/cm, this leads to a velocity of about 100 nm/s. Having in mind the natural diffusive motion with a diffusion constant of $D \approx 10^{-10}$ m^2/s[75,73], this directed electrophoretic motion is difficult to separate from the underlying fluctuations within the accessible time frame of simulations that usually is limited to microseconds.

This problem can be overcome by simulating long trajectories at the expense of computing time. Applying an artificially high external field also reduces the computational effort, but may induce a conformational change of the polyelectrolyte and its surrounding counterions, which leads to significant discrepancies between experimentally observed results and simulations. According to Netz *et al.* this effect is attributed to the polarisation and the following removal of the counterion cloud surrounding the polyelectrolyte[76,77]. Below a critical value, the mobility is not affected by the electric field, and the system is in the linear response regime.

Experimentally used electric fields are usually below 1 kV/cm, which corresponds to a reduced field strength of $E = 0.001$, and as such are far below the critical threshold. This study focuses on the behaviour of polyelectrolytes in weak electric fields below the threshold of $E \approx 0.2$. For simulations, the usage of a weak electrical driving force requires long simulations times, in order to be able to accurately separate the directed motion from the thermal fluctuations.

Alternatively, the electrophoretic mobility can be calculated from the following Green-Kubo relation

$$\mu = \frac{1}{3k_{\mathrm{B}}T} \sum_i q_i \int_0^\infty \langle \vec{v}_i(0) \cdot \vec{v}_{\mathrm{cm}}(\tau) \rangle \, d\tau, \tag{60}$$

where the summation is over all charged particles (monomers, counterions, and salt ions) in the system, and the $\vec{v}_i$ are their individual velocities and the q_i their charges. Here, $\vec{v}_{\mathrm{cm}}$ is the velocity of the center of mass of the polyelectrolyte. This approach has been successfully applied in simulations of charged colloids[78,8]. Please refer to the ph.d. thesis of Dr. Kai Grass[4] for a detailed derivation of Equation 60.

This method guarantees that no conformational changes of the chain structure or the ion distribution are induced by an artificially high external field. Another beneficial side effect of this method is that both transport properties can be obtained from the same simulation trajectories in the absence of an applied field without additional computational effort.

5 Electrophoresis of a Colloidal Sphere

In this section (with large sections taken verbatim from references[79,78,80]) we review work done on the electrophoresis of colloidal spheres using the hybrid LB-MD scheme. Instead of using a stationary exclusion represented by bounce-back boundary conditions to model the colloid, we use point particles ('monomers'). The coupling of the monomers which make up the colloid to the surrounding fluid is as described in Section 4.8.1. The flow velocity at the location of the monomer is obtained via linear interpolation from the surrounding lattice sites, and this implies that the lattice spacing should be of the order of the monomer size. The advantage of this approach is that it is quite flexible, because various large objects of soft matter physics (colloidal particles, polymer chains, membranes, etc) can be built up from elementary monomers, without major restructuring of the underlying simulation program. However, it is not possible to model the colloidal particle just in terms of a single monomer. While the electrostatic interactions plus the excluded volume between colloidal particle and counterions could be easily represented in terms of a single strongly charged particle with large repulsion radius, such a particle would have inappropriate hydrodynamic properties. The fluid would be coupled only to the centre of the particle, while a faithful representation of the hydrodynamics requires stick boundaries at the surface of the particle, or a good approximation thereof. Furthermore, we wish to faithfully represent the particle's rotational motion. For these reasons, we add a two-dimensional tethered network of monomers which we wrap around the surface of the central sphere, such that the overall structure resembles a raspberry as seen in Figure 8. The coupling to the solvent is then done only via the surface sites in terms of the monomer friction coefficient. For reasonably large friction, this is an excellent approximation to a

stick boundary condition, as has been shown in[81] in terms of both the translational and the rotational motion of the sphere, see also [82] .

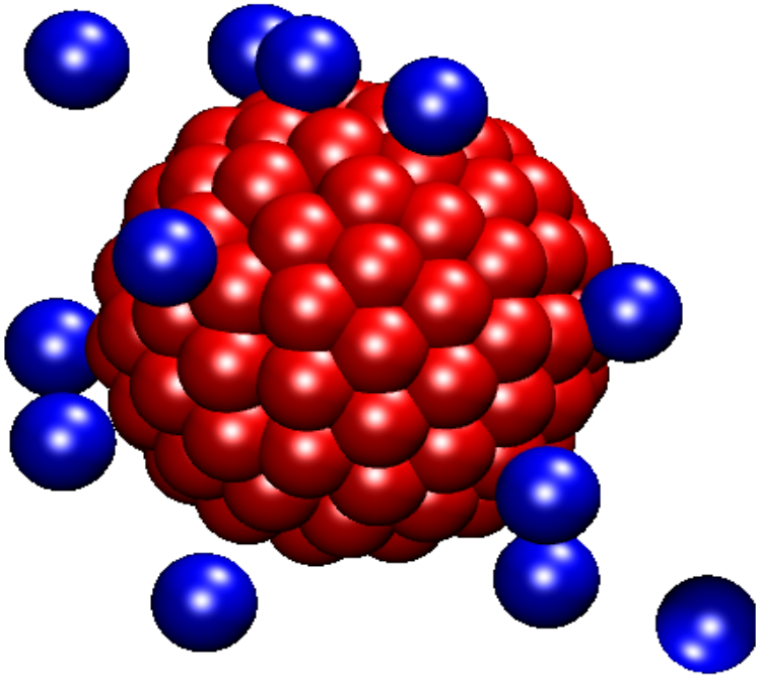


Figure 8: Raspberry-like model of a colloidal sphere. There is a central large bead of radius $R = 3$ and charge $Z = 20$. The small beads of radius 1 are connected with their nearest neighbors on the surface via FENE bonds. A repulsive soft-core potential is also operating between all the monomers. The counterions are moving freely in space and interact with the central bead via the Coulomb potential and the repulsive LJ potential. Figure taken from[79].

In the work presented here the electrophoresis of colloids of radius $R = 3$ was investigated using the aforementioned raspberry model. In order to define a dynamic effective charge of the colloid, which is related to the zeta potential, we look at the radial distribution of the ionic velocity in the double layer, taking the velocity component in the direction of the external field as shown in Figure 9. As the instantaneous velocities are governed by thermal fluctuations and thus have a very small signal-to-noise ratio, we calculated the average velocity over 10 LJ time units. This procedure averages out the stochastic components of the velocity so that the directional drift dominates the motion. The corresponding average curves are shown in figure 5. The velocity is shown relative to the mean velocity of the colloid. We see that the ions in the nearest surroundings of the particle surface move along with the colloid. The relative velocity reaches unity at about $r = 4$, which is in agreement with the shear plane being about one ion diameter from the surface. From there, the correlation decreases monotonically and turns negative at about $r = 8$. We then see a region of anti-correlation spanning until $r = 15$, the half-box distance. In this re-

gion, the ionic drift velocity is anti-parallel to the colloidal one. Since there is no well defined plateau near the particle, it is not possible to define a slip surface (where the ion motion stops being fully correlated with that of the central particle) unambiguously, although $r = 4.5$ (the dotted line) seems like a reasonable value. This defines, on the one hand, the effective hydrodynamic radius of the overall object, and on the other hand the effective charge (i.e. the charge within the slip surface). Furthermore, the zeta potential would be the electrostatic potential at this surface minus the potential at infinity.

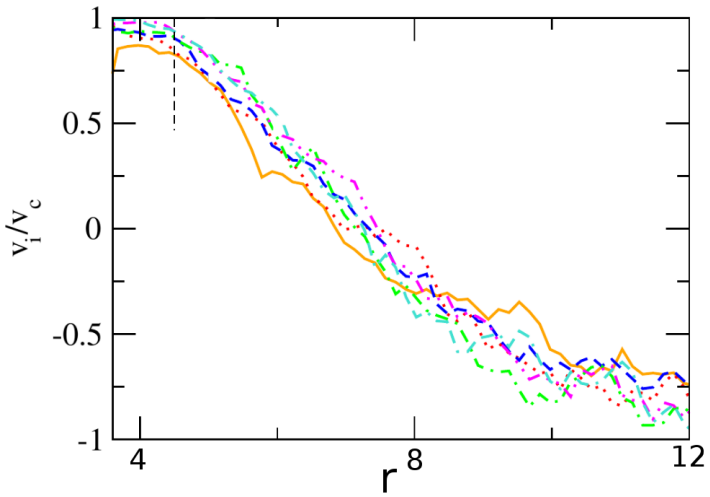


Figure 9: Radial distribution of the ionic drift velocity along the external field direction for colloids of different charge in an external electric field E = 0.2. The thin dashed curve indicates the approximate position of the slip surface. The various curves represent different bare colloid charges ranging from $Z = 40$ (the solid orange curve) to $Z = 120$ (the dash-dotted turquoise curve).Figure taken from[79].

In Figure 10 we examine the reduced mobility as a function of the bare colloidal charge Z. Three curves are shown, two where an electric field is applied to the system and the mobility is calculated directly as the velocity divided by the applied external field, and one using the Green-Kubo method as described in Section 4.12. While the direct method with an applied field of $E = 0.1$ is in good agreement with the mobility measure using the Green-Kubo method, the curve for $E = 0.2$ shows a significantly larger mobility. This is because the velocity is a nonlinear function of the applied electric field and mobilities calculated at higher electric fields tend to overestimate the colloid's mobility. It should be noted that the electric fields at which this occurs are on the order of 100,000 V/cm, which is much larger than what is achievable experimentally. The advantage of using a high electric field is an improved signal (velocity, which is proportional to E) to noise (diffusion, which is independent of E) ratio. For this reason most electrophoresis simulations try to use the highest electric field for which the velocity is still a linear function of the applied field The mobility itself first sees a linear increase in the mobility for low charges as one would naively expect. In this regime the electrostatic potential in the vicinity of the colloid's surface is relatively small and thus the nonlinear nature of the Poisson-Boltzmann

equation is not yet apparent. As the charge on the colloid increases the nonlinear terms in the Poisson-Boltzmann equation cause the counterions to strongly adsorb to the colloid's surface. These strongly adsorbed ions are also firmly hydrodynamically coupled to the colloid's surface due to their close proximity. When one examines the mobility values closely, one actually sees a slight drop in the electrophoretic mobility at the highest bare charges. This is only possible due to a second effect, often referred to as the relaxation force. The charged double layer is perturbed by the applied electric field, creating a dipole in the charge cloud. This perturbation results in an electrostatic force between the counterions and the oppositely charged colloid. This additional force on the colloid is in the opposite direction as that from the applied field, thereby reducing the colloid's mobility.

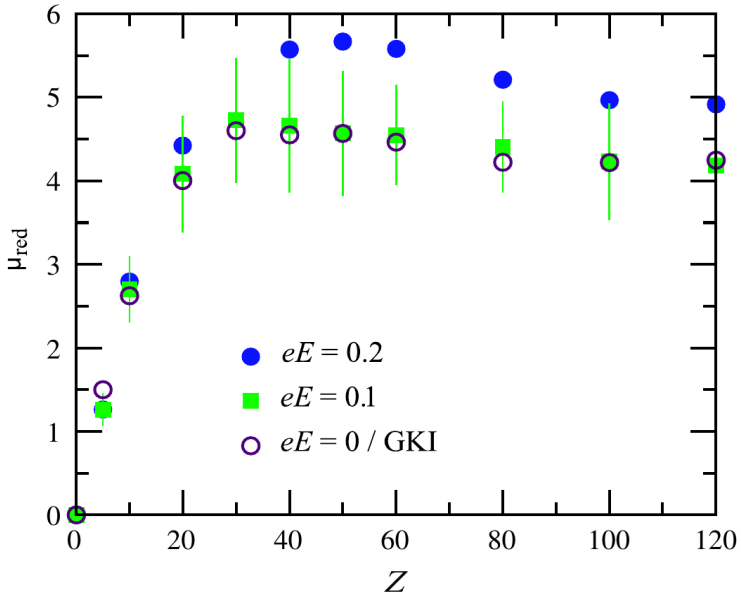


Figure 10: Reduced electrophoretic mobility of the colloidal particle, as a function of its charge Z . No salt is added, and, apart from the central colloid, the system comprises of Z monovalent counterions. The linear box size is $L = 30$. For strong driving field E, one observes nonlinear effects, while the results for weak driving agree favorably with the results of Green–Kubo integration (GKI). Note that the driving field is here given in the Lennard-Jones unit system of the simulation. Figure taken from[80].

Comparing the experimental and simulated mobility data in Figure 11, where the reduced mobility is shown as a function of the colloid density, one observes good agreement as long as the density is sufficiently large. However, simulation and experiment deviate in the regime of very low volume fractions. The reason is that at low densities the relative amount of the counterions in the experimental system becomes small compared to the back- ground ionic concentration (in this case, due to water dissociation), so that the dependence of the mobility on the volume fraction weakens. In the regime of background salt dominance, it should not depend on the colloid density at all, as it is assumed in the electrokinetic theories. In principle, one must expect that the salt species will introduce

yet another dimensionless parameter into the problem. However, we found that the effect of salt can be incorporated, within a reasonable approximation, by just adding the salt concentration to the counterion concentration, such that we obtain a new scaling variable κR with $\kappa^2 = 4l_{\mathrm{B}}(n_i + n_{\mathrm{salt}})$, where n_i is the counterion density and $n_{\mathrm{salt}})$ is the salt concentration. While not mathematically rigorous, this has been shown to be a reasonable approximation in closed nanoscopic systems over a wide range of salt and counterion concentrations[83]. The use of this procedure can be supported by the observation that the electrostatic potential in the regions centered between the colloids varies only weakly, such that a description in terms of a linearized PB equation is possible. In these regions, however, the local counterion concentration contributes to the screening parameter, too.

All of the simulations presented to this point were performed using only counterions. The unabsorbed counterions are then used to calculate an effective salt concentration. Figure 12 compares data obtained in a series of simulations at zero salt concentration and multiple colloids to results of simulations at constant but increasing amounts of added salt. The data coincide over a considerable region of R. The salt series curve continues with a slight increase, which is also a feature seen in the classical electrokinetic works[84,85]. Yet, we note that the minimum on the electrophoretic mobility can be related to the general behavior of the ionic double layer for highly charged colloids[86]. The agreement of the mobilities obtained for the same effective charge (total charge minus the number of adsorbed counterions) and R again supports our mapping postulate. Moreover, it clearly shows that the sort of ions present in the cloud is of minor importance for the mobility. The electrolyte effect, which leads to slowing down the particle drift, is produced in one case by solely counterions and in another case by both counterions and salt ions. In both cases it is only the total ion concentration that matters. Generally, we expect this scaling to be valid across the weak screening regime, $\kappa R \lesssim 1$, but deteriorate at high salt conditions, $\kappa R \gg 1$, where our assumptions of insensitivity to the ion parameters and the close correlation between the static effective surface potential and the zeta potential are no longer justified.

6 Free Solution Electrophoresis of Polyelectrolyte Chains

In this section we simulate the mobility of a single charged polymer (a polyelectrolyte) in bulk (free solution). The material in this section is largely taken verbatim from the thesis of K. Grass[87]. More details on polyelectrolyte electrophoresis can be found in the reviews[88,89].

We determine the electrophoretic mobility μ as the ratio between the measured center of mass velocity v_{PE} and the magnitude of the electric field E:

$$\mu = \frac{v}{E}.$$

For comparison, the results are normalized by the monomer mobility μ_1.

In a previous publication[90], we employed a mesoscopic coarse-grained model using molecular dynamics simulations in connection with a Lattice-Boltzmann (LB) algorithm to extend the theoretical understanding of polyelectrolyte electrophoresis on a more detailed level, and in particular, we intended to investigate the role of hydrodynamic interactions in these systems. Our results were able to match the free-solution electrophoretic

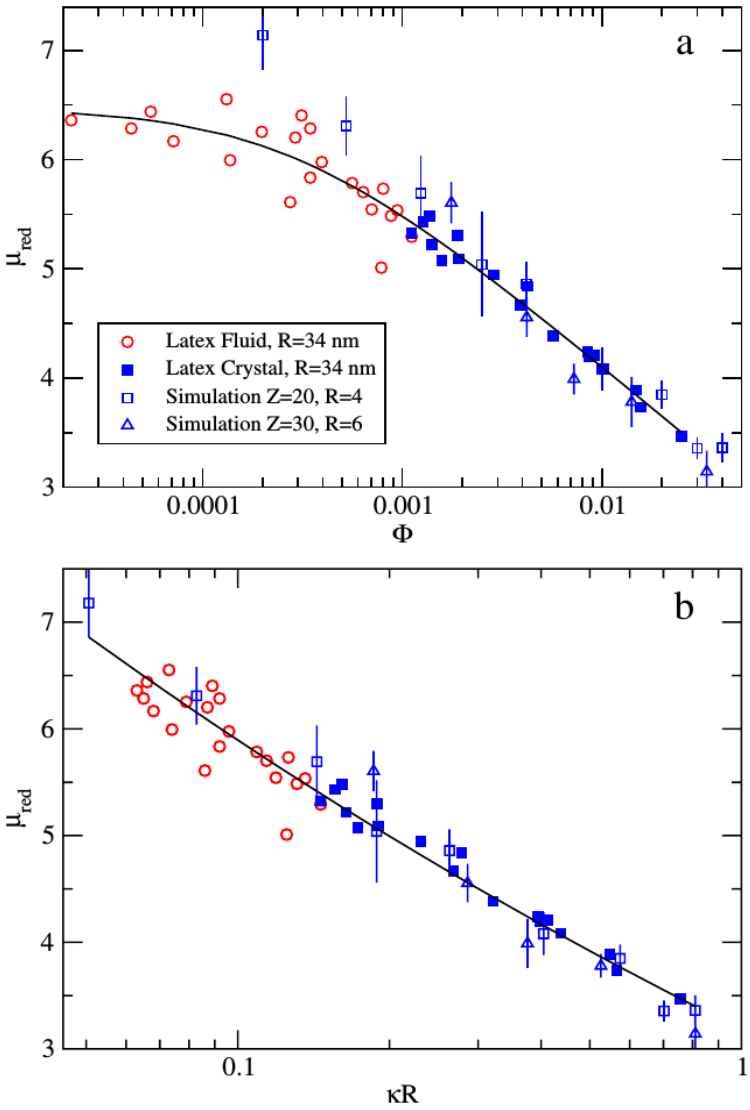


Figure 11: Reduced electrophoretic mobility of a spherical particle versus (a) particle volume fraction and (b) reduced screening parameter, calculated from counterion and salt contributions (see text) for various systems with matching Z. The solid curves represent a fit to the experimental data. Figure taken from[78].

mobility μ of short polyelectrolyte chains, here polystyrene sulfonate (PSS), as a function of the number of repeat units N with quantitative agreement to experiments as shown in Figure 13. Since the three data sets have different solvent viscosities the mobility is normalized by the corresponding constant mobility for long chains, the so-called free-draining mobility, μ_{FD}. The electrophoretic mobility increases for short oligomers, reaches a maximum for intermediate degrees of polymerization, and slowly decreases towards a plateau

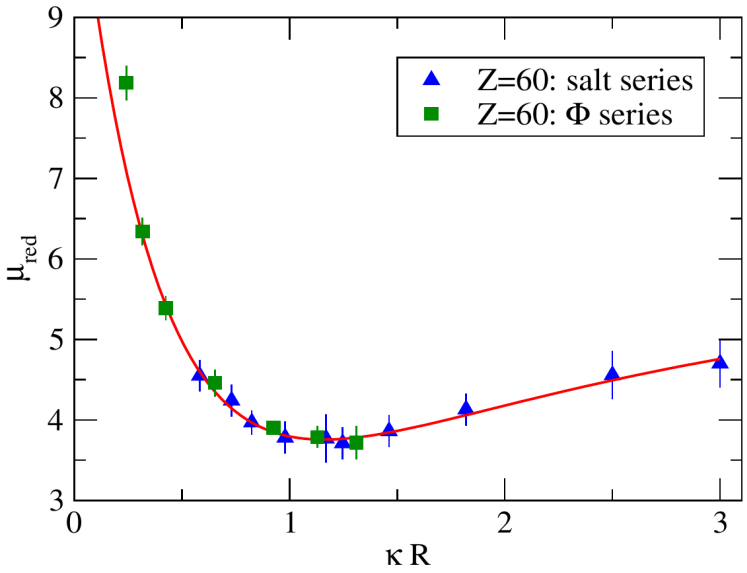


Figure 12: Variation of the electrophoretic mobility of particles of charge $Z = 60$ in an external field $E = 0.1$ versus the reduced screening parameter. The screening parameter was calculated from the concentration of the free ions at different colloid volume fractions and/or concentrations of added 1:1 salt. The curve is a guide to the eye. Figure taken from[78].

value for long chains. To understand this observation, the hydrodynamic interactions were investigated in detail and we found that they are actually the major driving force for the length dependent mobility for short and intermediate chain lengths. The constant mobility for long chains can be attributed to an effective screening of hydrodynamic interactions, which leads to the so-called free-draining behavior. The inset in Figure 13 shows a comparison to a coarse-grained simulation that neglects hydrodynamic interactions. This leads to a qualitatively completely different behavior, showing a monotonically decreasing mobility. Agreement to the experimentally observed behaviour is only achieved as long as hydrodynamic interactions are included correctly as has been shown in detail in our previous investigations[91,90].

Figure 14 displays the characteristic behaviour of flexible polyelectrolytes for vanishing salt concentration $c_s = 0\,\mathrm{mM}$: initially, the electrophoretic mobility increases with N to reach a maximum at intermediate chain lengths and then slowly decays towards a constant value for long chains. This constant value, often called the free-draining limit μ_{FD}, can be explained by the length independence of the ratio between effective charge and effective friction for long chains as we will show later.

In the presence of added salt, the long chain mobility is reduced, which is consistent with the experimentally observed behavior[92]. Furthermore, the shape of the curve is influenced, and the maximum at intermediate chains is suppressed for increased salt concentration. At $c_s = 160\,\mathrm{mM}$ the maximum disappears and the measured mobility becomes length independent within the resolution of the simulation. A further increase of the added salt concentration leads to a further reduction of the limiting mobility μ_{FD}, not shown here,

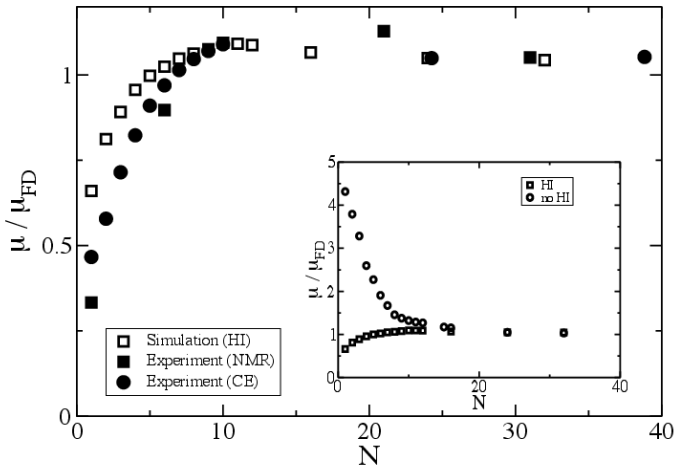


Figure 13: The normalized electrophoretic mobility $\mu/\mu_{\rm FD}$ as a function of the number of repeat units N for simulation data including hydrodynamic interactions (HI), and experimental data coming from capillary electrophoresis (CE) and from electrophoretic NMR. The inset compares to simulation data obtained with a model neglecting hydrodynamic interactions. Figure reproduced from[87].

while the monomer mobility μ_1 remains almost unchanged. This leads eventually to an inverted length-dependence with a monotonic decrease of the mobility towards the limiting value.

In the simple local force picture, the constant center of mass velocity $v_{\rm PE}$ that determines the electrophoretic mobility is a direct result of the cancellation of two acting forces: the electric driving force $F_{\rm E} = Q_{\rm eff}E$ is canceled by the solvent friction or drag force $F_{\rm D} = \Gamma_{\rm eff}v_{\rm PE}$. Here, $Q_{\rm eff}$ is the effective charge of the polyelectrolyte, which can be thought of as the bare charge of the polyelectrolyte reduced by oppositely charged ions in solution that associate to the polyelectrolyte chain. The association of counterions to a PE chain is known as counterion condensation[93,94]. The compound formed by the polyelectrolyte and the associated ions is moved through the solvent under the influence of the external field and experiences a Stokesian drag force with an effective friction coefficient $\Gamma_{\rm eff}$ that is a priori unknown. In the steady state both forces balance and the mobility is given by

$$\mu = \frac{v}{E} = \frac{Q_{\rm eff}}{\Gamma_{\rm eff}}.$$

Next, let us compare the results of Figure 14 to the case when long-range hydrodynamic interactions between the particles are neglected in simulations, *i.e.*, by using a standard Langevin thermostat. The results are shown in Figure 15. One immediately notices that the observed electrophoretic mobility differs significantly from the behaviour observed in Figure 14. Independent of the salt concentration, the mobility decreases monotonically

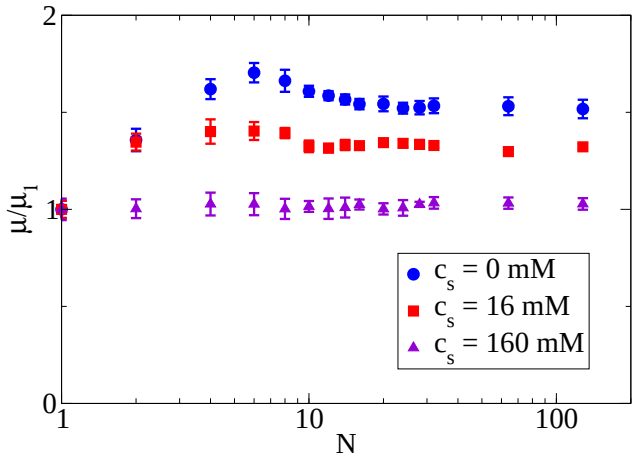


Figure 14: The normalized electrophoretic mobility μ/μ_1 of polyelectrolyte chains of length N for three different salt concentrations using the LB algorithm. The added salt not only influences the absolute mobility, but likewise changes the characteristic shape of the mobility with respect to chain length N. Figure reproduced from[87].

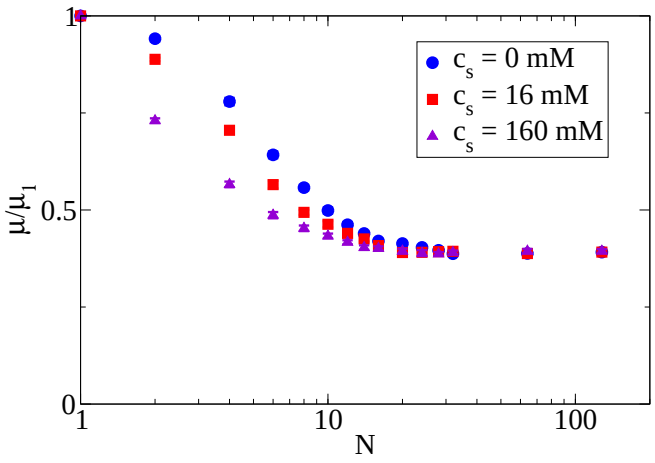


Figure 15: The normalized electrophoretic mobility μ/μ_1 for different chain length N at varying salt concentrations c_s without hydrodynamic interactions differs significantly from the behaviour observed in Figure 14. The mobility shows a salt-dependent monotonic decrease for short chains and a salt-independent constant value for long chains. Figure reproduced from[87].

with chain length and slowly approaches a constant value for long chains which is independent of the salt concentration. While the mobility is indeed chain length independent for

long chains, the qualitative behaviour is drastically different compared to both the experimental results and simulations including hydrodynamic interactions. This is an excellent example of how failure to account for hydrodynamic interactions can fail to even qualitatively reproduce the physics of a system.

7 Electrophoresis of a Polyelectrolyte-Grafted Colloid

In this section we examine the mobility of a polyelectrolyte-grafted colloid, the material for this section is largely taken verbatim from[95]. The electrophoresis of bare charged particles of different shapes and surface charges, such as DNA, spherical colloids, and flat surfaces has been extensively studied in the past decades. Recently, specific attention was drawn to the phenomenon of mobility reversal of such particles as a function of the ionic strength. This mobility reversal is, however, related to charge inversion and thus, limited to multivalent counterions and relatively high surface charge densities and salt concentrations[96–106]. The charge inversion is caused by the strong correlations of the multivalent counterions and as such does not occur with monovalent salt ions.

The electrokinetic properties of polyelectrolyte-coated spheres, usually referred to as soft particles, are significantly more complex than that of bare particles due to the nonuniform nature of the surface charge distribution and the additional hydrodynamic drag of the polymers[107–111]. It is necessary to take these additional effects into account correctly to model the electrophoretic mobilities of biological cells which often have naturally occurring polymer layer coatings on their surface[112–117]. In addition, artificial polymer coatings are used to control the electrokinetic properties of surfaces[118–120], and the rheological properties of electrorheological fluids[121,122]. Polymer-grafted colloids also show great promise as drug and gene delivery vehicles[123–125]. The broad range of applications has generated extensive theoretical, numerical, and experimental work on the electrophoretic properties of diffuse, soft particles[126,109,127–131,115,132].

In a recent publication[95], we reported on a molecular dynamics (MD) simulation study of the electrophoresis of negatively charged colloids grafted with positively charged polymers in the presence of monovalent salt of varying concentration. The simulation were performed using the ESPResSo package[133,134] and all particles were considered explicitly, as shown in Fig. 16. Full hydrodynamic interactions were taken into account using a GPU-based Lattice-Boltzmann (LB) algorithm[3]. We investigated the mobility for two distinct cases; in the first case the soft colloid as a whole was charge neutral, i.e. the colloid charge was exactly balanced by the charges on the grafted polyelectrolytes. In the second case, the magnitude of the colloidal core's charge was larger than the charge on the grafted layer and thus the soft colloid had a net negative charge of $Q_{\mathrm{net}} = -50$ e.

Fig. 17 shows the results for the mobility as a function of salt concentration for both cases. We found that the net-neutral soft colloid had a positive mobility at moderate and high salt concentrations. The negatively net-charged soft colloid reversed mobility as a function of *monovalent* salt concentration. The physical origin of this mobility reversal is different than that of the bare surfaces as it does not rely on the strong correlations of multivalent salt. A careful examination of the radial monomer and ion density profiles and comparing the two important length scales of the system, namely the Debye length λ_{D} and the polymer layer's thickness H, revealed the mechanisms causing these behaviors.

Let us first examine the low salt limit of the mobility. In this regime, the mobility is zero

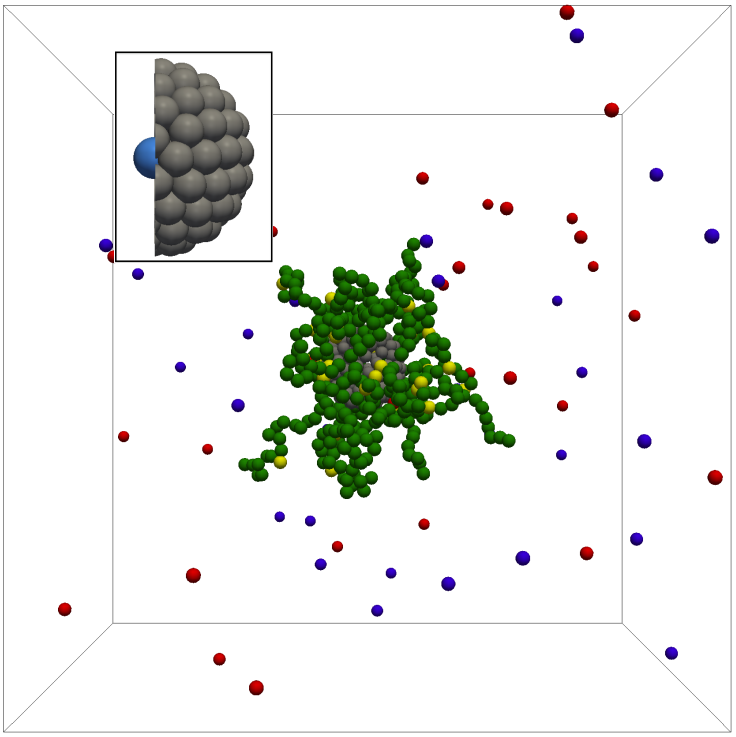


Figure 16: A simulation snapshot of the neutral soft colloid at $c_s = 0.01$ M showing the neutral monomers (green beads), charged monomers (yellow beads), positive salt ions (red beads), negative salt ions (blue beads), and the surface raspberry beads (grey beads). The inset shows a cut of the raspberry with the negative central particle in blue. Figure taken from[95].

for the net-neutral soft colloid and negative for the net-negative soft colloid, as could be expected. At low salt concentrations there are hardly any salt ions in the system, making the Debye length longer than the characteristic dimensions of the colloid. In the Hückel limit of no salt, the mobility of a spherical particle is given by $\mu = Q_{\mathrm{net}}/(6\pi\eta R_{\mathrm{H}})$, where R_{H} is the hydrodynamic radius of the object. While this explains the results for the net-neutral case, using this relation for our net-charged soft colloid with $R_{\mathrm{H}} = R_{\mathrm{col}} + 2R_{\mathrm{G}}^{\perp} \sim 7\sigma$ gives $\mu_{\mathrm{red}} \approx -15$ which is much larger than the simulation mobility at $c_s = 0.001$ M. $R_{\mathrm{G}}^{\perp}$ is the component of the radius of gyration perpendicular to the core's surface and has a value of $R_{\mathrm{G}}^{\perp} \sim 2\sigma$. $2R_{\mathrm{G}}^{\perp}$ is a measure of the brush thickness. The high density of counterions near the colloid's surface seen in Fig. 18 (d) shows that the system is still far from the Hückel limit. Due to the finite colloidal concentration, there is still have a significant amount of counterions around the colloid which reduce its bare charge, and thus lower the mobility[136]. Besides, the polymer layer hinders the motion of the ions and the fluid, thus creating an additional drag force.

Upon increasing the salt concentration, λ_{D} progressively shrinks and more ions penetrate the brush. The electroosmotic flow (EOF) of these ions is non-uniformly dampened by

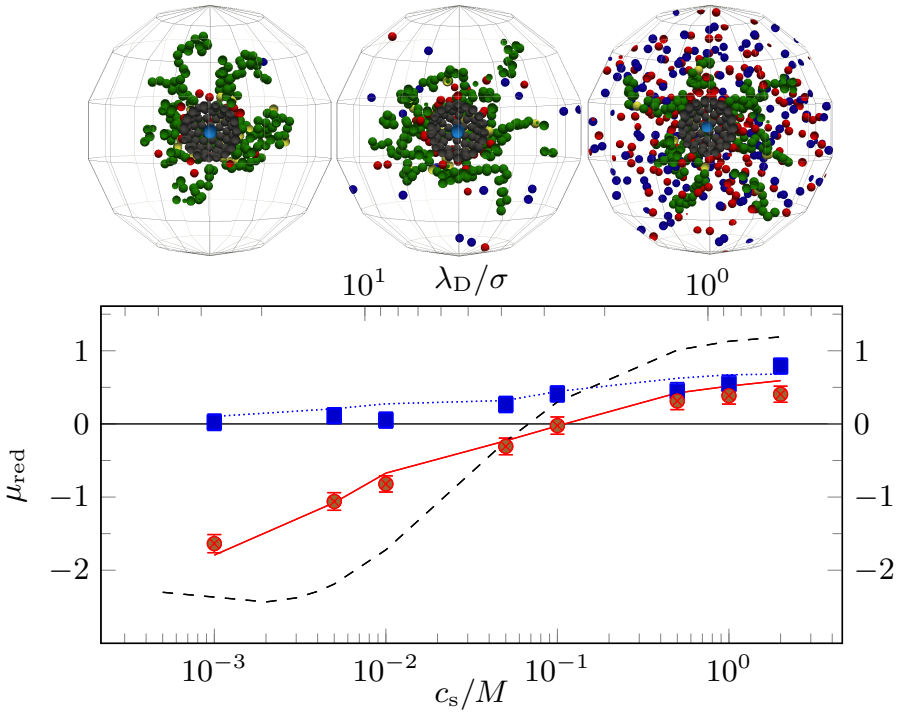


Figure 17: The reduced mobility μ_{red} as a function of salt concentration c_{s} and the corresponding Debye length λ_{D}. MD simulations for a neutral soft colloid with $Q_{\mathrm{net}} = 0$ e (blue circles) and a charged soft colloid with $Q_{\mathrm{net}} = -50$ e (red circles) are shown. Numerical results are depicted by the solid lines (blue for $Q_{\mathrm{net}} = 0$ e and red for $Q_{\mathrm{net}} = -50$ e). The dashed line shows the numerical result for a negatively charged soft colloid with the same colloid and net charge density as used for the red circles, but with $R_{\mathrm{col}} = 1$ μm. Figure taken from[95].

the brush, the closer to the colloidal core's surface, the stronger the screening. Therefore, the EOF generated by the core's counterions is mostly screened, whereas the competing contribution of the oppositely charged polyelectrolytes to the EOF partially reaches the bulk. The accumulation of ions within the brush is most pronounced when λ_{D} is commensurate with H. This is why the net-neutral soft colloid acquires a positive mobility at about $c_m athrms = 0.05$ M ($\lambda_{\mathrm{D}} \simeq 4\sigma$). Of particular note is that the mobility of the negatively charged soft colloid changes sign around $c_{\mathrm{s}} = 0.1$ M. This salt concentration corresponds to $\lambda_{\mathrm{D}} \sim 3\sigma$ which is slightly smaller than H. In Fig. 18, we see that the accumulation of counterions is significantly greater at $c_{\mathrm{s}} = 0.1$ M compared to $c_{\mathrm{s}} = 0.001$ M. At the highest salt concentration, $c_{\mathrm{s}} = 2.0$ M, the mobilities in Fig. 17 saturate at a finite value, in contrast to bare colloids where the mobility is negligible at high ionic strengths[137,138]. The mobilities are roughly the same regardless of the colloidal core's charge. This happens because the Debye layer is so thin that sufficient counterions aggregate within a few bead diameters of the colloid surface to neutralize it, as seen in Figs. 18(b and d). Since the polymer brush is rather thick, the EOF generated by these ions is almost completely screened by the polymers and the mobility is largely determined by the properties of the

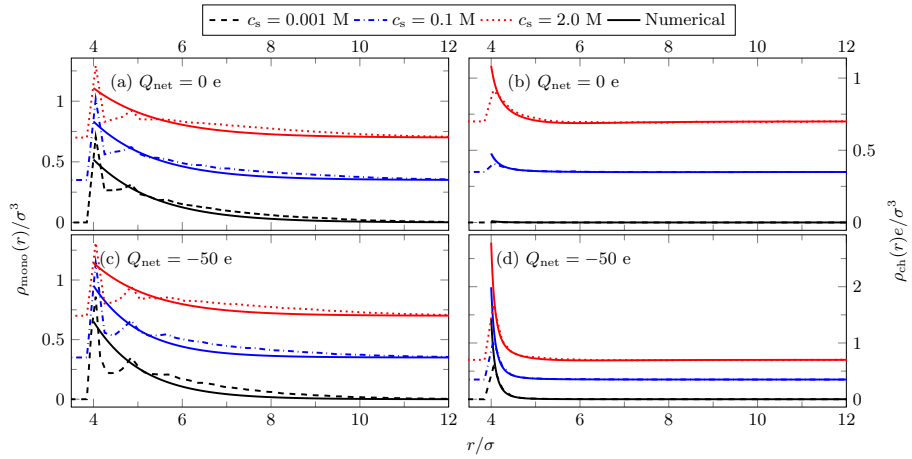


Figure 18: Radial density profiles for $c_s = 0.001$ M (dashed black lines), $c_s = 0.1$ M (dashdotted blue lines) and $c_s = 2$ M (dotted red lines) of the monomers $\rho_{mono}(r)$ and fluid charge, i.e. the sum of the ions' charges excluding the charges fixed on the polymers, $\rho_{ch}(r)$. Panels (a) and (b) show the data for the neutral soft colloid while panels (c) and (d) show the results for $Q_{net} = -50$ e. All curves are shifted vertically by factors of 0.35 for the sake of visibility. The solid lines represent the values produced by the program from Hill *et al.*[135]. Figure taken from[95].

coating polyelectrolytes[119,120,135,139]. This is in line with theoretical predictions for the electrophoretic mobility of soft surfaces at hight salt concentrations. According to the analytical expression derived by Ohshima *et al.*[126], in the limit of $\kappa \to \infty$ the mobility is proportional to the charge density of the grafted polyelectrolyte layer. This results in $\mu_{red} \sim 0.6$ for our parameters, which agrees with our simulation and numerical results.

It is important to note that the nature of the mobility reversal that we observed for the net-charged soft colloid with respect to *monovalent* salt concentration is completely different than that observed in the case of bare colloids in the presence of multivalent salt. Whereas the latter is caused by the overcharging of the colloid due to ion correlations, the former is a mere result of the interplay between the EOF and the screening effect of the brush.

We also compared all of our mobility results in Fig. 17 to numerical results of the extended electrokinetic model of Hill et al.[135] and found good agreement. In this numerical approach, the electrokinetic equations are based on the Darcy-Brinkman formalism and solved in a manner similar to the standard electrokinetic model[84,85]. The qualitative trend as a function of salt concentration were the same as in the simulations for $Q_{net} = 0$ e and $Q_{net} = -50$ e. In both cases the slight differences between the numerical and simulation results were mostly due to the fits to the simulations' monomer densities shown in Figs. 18(a and b), which were used as an input into the numerical calculations. The good agreement of our simulation results with the numerical approach confirmed the validity and applicability of the Darcy-Brinkman equation to describe electrokinetic phenomena of soft particles.

Using the numerical solver of Hill et al., we further showed that the reversal also occurs for experimentally realizable colloids at experimentally accessible salt concentrations.

This is shown by the black dashed line in Fig. 17. The salt concentration at which the mobility reverses sign is almost the same as the one in the case of the smaller colloid since the mechanism causing it is independent of the colloid size and only depends on the ratio of the two important system length scales, namely, λ_{D} and the thickness of the polyelectrolyte layer.

8 Nanopore Conductance

8.1 Lattice-Boltzmann/MD simulations

In this section we will present a model of ion transport through nanopores based on coarse-grained simulations. This section is copied verbatim from[66] with some minor modifications. The modeling strategy relies on treating all charged objects as explicit particles, and replacing the water molecules by a continuum. As dynamic quantities are considered, a simple dielectric picture is not sufficient. The water, like any solvent, assumes an active role in the dynamics. The two main effects of water are the electroosmotic flow and hydrodynamic interactions. Electroosmotic flow occurs as a collective phenomenon, where the concerted motion of many ions induces water flow. Hydrodynamic interactions are also relevant in ion conduction phenomena. We will show below that they cause, for example, a decrease of the specific conductivity of an electrolyte with increasing concentration.

For the systems consisting of ions and water, the hydrodynamic radius of the particles must be matched. This is the key prerequisite to obtain the correct ratio of convective and direct ion transport and therefore must be matched in our mesoscopic approach. The numerical value of the viscosity, on the contrary, is not important. By using appropriate scaling factors we can a posteriori deduce the physical values of all observables. We apply the LB model to the problem of nanopore conduction using an infinite cylinder. We systematically compare the results of the Lattice-Boltzmann simulations to those produced by the standard electrokinetic equations in an infinite cylinder.

8.2 Nanopore conduction with the LB model

8.2.1 LB simulations of the conductivity of a nanopore

The hybrid LB-MD scheme outlined in Section 4 was applied to calculate the conductance of a nanopore in the presence and in the absence of a coarse-grained double stranded (ds) DNA molecule. Inspired by the electrokinetic continuum model, the DNA molecule was represented as a charged cylinder and the nanopore as a cylindrical channel. We considered a system where both the pore and the DNA are infinite based on periodic boundary conditions along the channel. This way we can simulate a finite piece of pore and DNA that imitates an infinite system. The DNA cylinder was charged by placing particles with one elementary charge at a mutual distance of 0.16 nm along the cylinder axis. The DNA-ion interaction was modeled by a WCA interaction with $\epsilon = k_{\mathrm{B}}T$ and σ =0.425 nm, and the distance was shifted by 0.575 nm to obtain a steep potential and a distance of closest approach of around 1 nm. The Bjerrum length of 0.71 nm reflects a temperature of 298 K and an aqueous solvent. A box length of 25 nm was chosen and again all lengths were expressed in units of the LB lattice constant of 0.34 nm.

The total charge of the cylinder consisted of 144 elementary charges. The system was neutralized by adding 144 positive counterions, and a varying number of extra ions was added to the system. In this geometry it is impossible to use an ion reservoir. Our strategy was therefore adding a different numbers of salt ions, perform the simulation and determine the concentration of a hypothetically attached reservoir a posteriori. This leads to a somewhat different salt concentration than one obtains by dividing the number of ions by the pore volume. We determined the salt concentrations by calculating the local concentration as a function of the distance to the pore axis, and fitting a constant to the mean of the concentration of positive and negative ions in a region between 3 and 4.5 nm away from the pore axis. If the Poisson-Boltzmann equation held exactly, this method would yield the exact same results given that the electrostatic potentials are smaller than $k_{\mathrm{B}}T/e$. This is the case in the regions used for averaging. The same method was also applied for the atomistic simulations described below, where we performed further simulations to justify it.

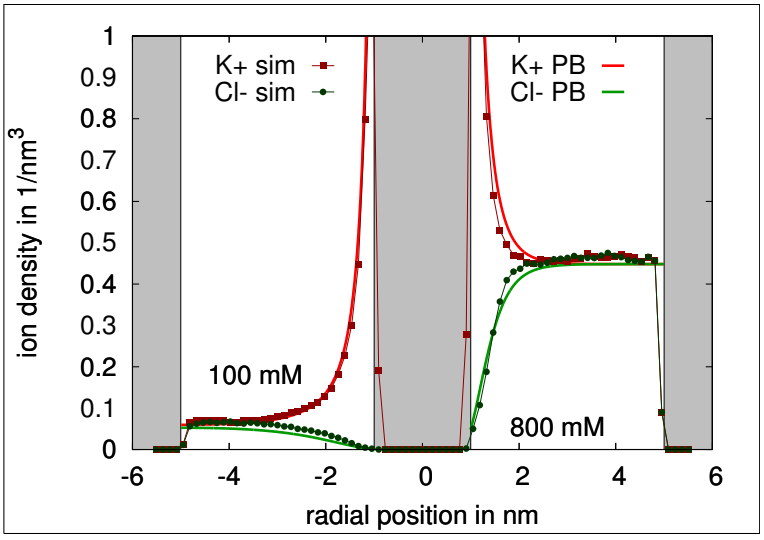


Figure 19: Radial density profile of ions. K^+ ions are shown as squares, and Cl^- ions as circles as obtained from the simulations. Corresponding results of the electrostatic continuum model are shown as lines. The left half of the plot is corresponds to a reservoir concentration of 135 mM and the right half to 600 mM.

In Fig. 19 we display the ion densities we obtained from our simulations for two different data sets, corresponding to 101 and 505 ion pairs added to the system. The corresponding reservoir concentrations were determined to be 135 mM and 600 mM. The agreement between the electrokinetic infinite cylinder model and the simulations is good. Especially for the higher salt concentration case the decay of the counterion cloud towards the bulk value is somewhat faster in the simulations compared to the PB theory. This indicates that some ion correlation effects are visible that are present in our coarse-grained simulations and not contained in the mean-field PB theory. Overall, however, the PB theory does a good job reproducing the results of the coarse-grained simulations. The usage of the LB

method here is irrelevant since hydrodynamic interactions do not affect the static ion density distributions. We therefore can say, consistent with previous findings[140], that the ion density and hence also the electrostatic potential around a charged rod with the line charge density of DNA is described well by the Poisson-Boltzmann equation, and that correlation effects do not play a major role.

Next, we want to examine the velocities of the ions and fluid as a function of radial distance. In order to do this we use a slightly modified version of the standard electrokinetic model using a position and concentration dependent mobility for the ions given by:

$$\mu\left(r\right) = \mu_0 \times f_{\text{conc}} \times \left(1 - \frac{8}{9}\frac{a}{r - r_{\text{DNA}}}\right) \times \left(1 - \frac{8}{9}\frac{a}{r - r_{\text{wall}}}\right), \tag{61}$$

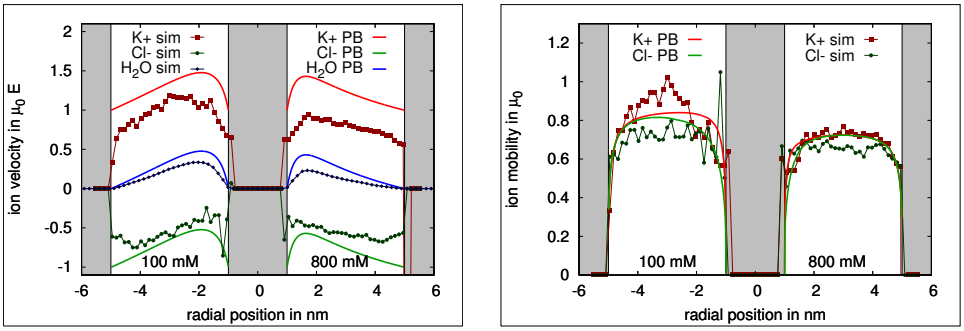
where f_{conc} is a factor based on experimental conduction data which takes into account the concentration dependence of the ions' mobilities. We use the concentration of ions of opposite sign when calculating f_{conc}, since the reduction in ionic mobilities due to retardation forces and collisions with other ions are mostly the result of interactions with ions of opposite charge. The terms in parenthesis account for the increased friction of the ions as they are in the vicinity of the DNA at r_{DNA} or the wall at r_{wall} and are based on the references[141,142]. A more detailed description of this modification to the standard electrokinetic model can be found in the Doctoral thesis of Stefan Kesselheim[66].

The position-dependent water and ion velocities from this modified version of the standard electrokinetic model, together with our LB simulation results and the experimental data of Smeets *et al.* [143] are depicted in Fig. 20a. The water velocities obtained in the simulations are larger in the simulations than in the electrokinetic model. This is true for both data sets shown, although the deviations for higher salt concentrations are larger. Qualitatively the fast decay of the counterion cloud in the simulations will lead to a smaller water flow. This can be understood by considering the extreme case of ions sitting right at the boundary. The force acting on them is immediately transferred into the DNA and they do not produce electroosmotic flow. The effect can be considered analogous to a lever. The larger the distance of the counterions from the DNA, the larger is the EOF they create. This qualitative finding is also the reason why the EOF is smaller when the Debye length gets smaller. A second possible cause is related to the coupling scheme of the LB particles. Due to the multilinear assignment scheme forces in a certain range are directly transferred into the DNA without causing an acceleration of the water. This could be a second source of deviation, which we have not completely ruled out. Finally it is interesting to note that the position of the hydrodynamic boundary is found exactly at the desired distance from the pore axis, although the cylinder is represented on a lattice. The ion velocities we obtain are in good qualitative agreement with the values obtained from the continuum model, although significantly smaller. This is not surprising as the conductance decreases with increasing concentration in the coarse-grained simulation model.

In the electrokinetic continuum model the ion velocity is just the water velocity plus a constant offset. Inspired by this notion we define an effective position-dependent mobility as the difference between the local ion velocity and the water velocity, as

$$\mu^{\pm} = \pm\left(v_z^{\pm}\left(r\right) - u_z^{\pm}\left(r\right)\right). \tag{62}$$

It is shown in Fig. 20b for both ion species, again for the same two data sets as used before. We normalize their values by the mobility of an isolated particle. At least for



(a) (b)

Figure 20: (a) Ion velocity as a function of the distance to pore axis. K^+ and Cl^- ions as squares and circles resp. The velocity of the LB fluid replacing the water is displayed as diamonds. The corresponding continuum model prediction are shown as lines. (b) Ion mobility as a function of the distance from the pore axis pore axis. Figure reproduced from[66].

the higher concentration data set the mobility is almost constant in the center between the DNA and the pore wall. This constant, however, is significantly smaller than unity, which reflects the concentration dependence of the conductance. Towards the walls the mobility decreases significantly. This observation is discussed in more detail below. The mobility of counterions is somewhat larger than the mobility of the coions. For this behavior we have so far not found a satisfactory explanation, but it is observed throughout our data.

Next let us examine the modulation of the total current in the pore, which we define as

$$\Delta I/I = \left(I_{\text{with DNA}} - I_{\text{without DNA}}\right)/I_{\text{without DNA}}. \tag{63}$$

The results are shown in Fig. 21 for the experiments, the LB simulations, and the standard electrokinetic model (assuming that the mobility of the ions is constant). The standard electrokinetic always shows an enhanced current for the salt concentrations shown, and the predicted current is always higher than both the experimental results and the LB simulations. This is not surprising since it neglects both the increased friction experienced by ions in the vicinity of the wall and the reduced mobility of ions at higher salt concentrations. When one uses the modified mobility tensor in Eq. 61 the agreement improves, but the crossover from current enhancement to current blockade still only occurs at 800 mM. The LB simulations do even better, predicting a crossover at around 700 mM. The slight difference between the two can be understood by the slight differences in the position dependent mobilities and the electroosmotic flow seen in Fig. 20. While the modifications of the standard electrokinetic model reproduce the LB results, there is still a large discrepancy in comparison to the experimental results of Smeets *et al.* [143].

Clearly both the standard electrokinetic model and the LB simulations fail to quantitatively reproduce the experimental results. We have, however, identified four important contributions to the current inside the pore, (a) the excluded volume of the DNA which ions can not enter, (b) the counterion cloud of the DNA which brings extra ions into the pore, (c) the electroosmotic flow and (d) the mobility reduction near boundaries. The exact shape, position of the charges, surface structure and other microscopic details determine

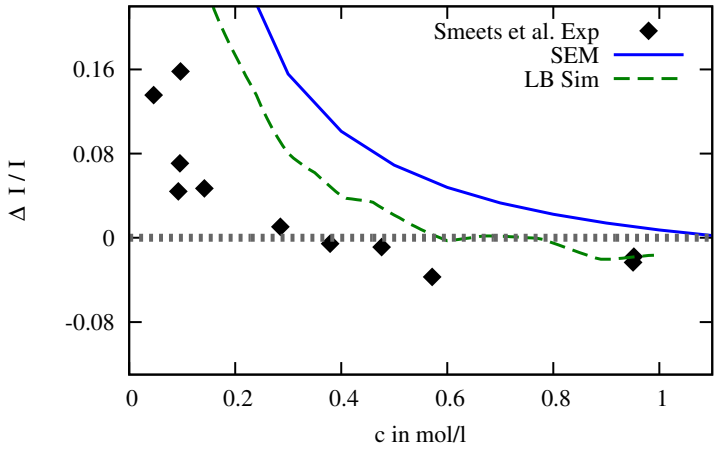


Figure 21: The current modulation as a function of the ionic strength. The experimental data of Smeets *et al.* [143] is shown as black diamonds, which shows a larger deviation both compared to the LB simulations (dashed green line) and the standard electrokinetic model (blue circles).

the relative magnitude of these effects. For this reason we decided to perform atomistic simulations as described in the next section.

8.3 Atomistic model of nanopore conductance

In this section we present our atomistic resolution model of a DNA in a nanopore. From the previous models we have learned what are the key effects that determine how much the nanopore current is modulated by the presence of a DNA molecule: (a) the excluded volume of the DNA, (b) the DNA charge and the structure of the counterion cloud (c) electroosmotic flow, and (d) friction effects near the DNA. The magnitude of each effect however remains unclear. All of them are significantly affected by the microstructure of the DNA. Therefore we studied these quantities with a model with atomic resolution to achieve a more realistic picture.

The guiding principle of atomistic simulations is that every atom is represented by a particle. While computationally quite costly, atomistic simulations are more realistic and offer insights which cannot be gained through coarse-grained approaches. The possibility to zoom in on details which are experimentally inaccessible allows scientists to gain an entirely new perspective on microscopic mechanisms, that, for example, are *e.g.*, biologically relevant.

In Fig. 22 we outline our simulation setup. Just as in our previous simulations, we investigate the central part of a cylindrical nanopore. The system length is chosen in agreement with the symmetry of a DNA molecule. We use a simulation box size of 6.76 nm, which contains exactly two full turns of the double helix. We apply periodic boundary conditions to mimic an infinite system. We choose a DNA homopolymer which consists only of CG base pairs. The DNA is held in place by harmonic springs attached to the phosphorous atoms of the backbone, so that it can deform and fluctuate but cannot move away

from its initial position. Water and ions are added to the system. An electric field is applied parallel to the pore axis and the ion motion is monitored and analyzed. For comparison, two simulation setups with and without a DNA molecule were performed. In principle, this system mimics as close as possible the continuum model based on the standard electrokinetic equations and the hybrid LB-MD model. The only difference is that more classical degrees of freedom are taken into account explicitly in terms of particles. Similar simulation studies have been conducted by Luan et. al.[144], but with a focus on measuring the force acting on the DNA.

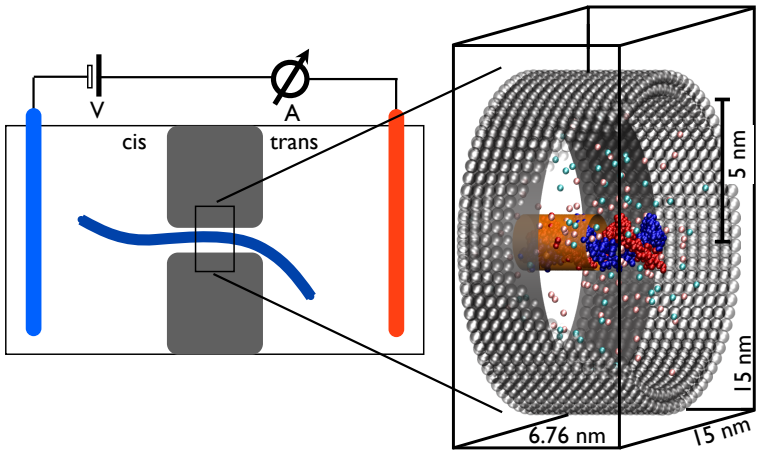


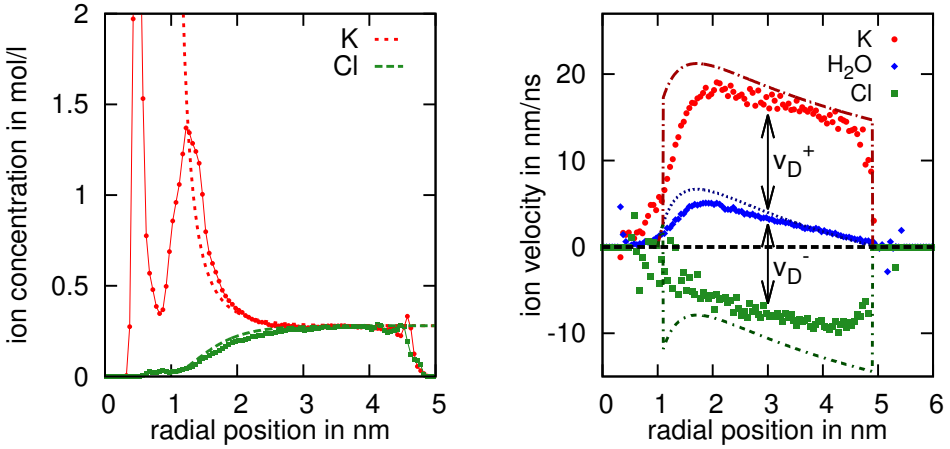
Figure 22: The simulation setup used for the atomistic simulations. We focus on the central, cylindrical part of a long nanopore. 20 base pairs, corresponding to two full turns of the helix are simulated in periodic boundary conditions. The pore is formed by Lennard-Jones bead forming a cylinder of radius 5 nm. A box of $6.76 \times 15 \times 15\,\mathrm{nm}^3$ is used, and an electric field is applied along the pore axis.

For the simulations, we used the MD package GROMACS 4.5.5[145,142,146]. The somewhat difficult choice of a combination of a force field, a water model, and interaction parameters of KCl are discussed in the original publications[147,66]

8.4 Atomistic simulation results

Our goal is to perform a detailed investigation of the static and dynamic properties of the system. It is split into two parts: First, the density profiles $c^{\pm}(r)$ and velocity profiles $v^{\pm}(r), u(r)$ are compared to the continuum model. Second we determine the electric current through the pore and compare this to the experimental results of[143].

The static ion density profile reflects the typical behavior known from the literature (*e.g.*,[148,144,149]). Fig. 23 shows, as an example, the simulation with 64 added ion pairs. The counterion density exhibits two significant peaks. One peak occurs at r=0.5 nm (full height not shown), and a second peak is observed at r=1.25 nm. The first peak is related to counterions entering into the major and minor grove of the DNA. This peak compensates for around 1/3 of the DNA's bare charge at all considered salt concentration. The second peak indicates the effective radius of the DNA backbone. For larger distances than 1.25 nm,



(a) (b)

Figure 23: (a) Comparison of the ion concentration as a function of the distance r from the pore axis at around 0.3 M of the atomistic data (points) to the continuum model (broken lines). The long distance decay of the second peak at $r = 1.25\,\mathrm{nm}$ is well described by the continuum model. (b) Comparison of the measured ion and water velocities as a function of r.

the ion distribution is well described by the continuum model. In the region between $r = 3\,\mathrm{nm}$ and $r = 4\,\mathrm{nm}$ an almost complete screening of the DNA charge was observed for all salt concentrations.

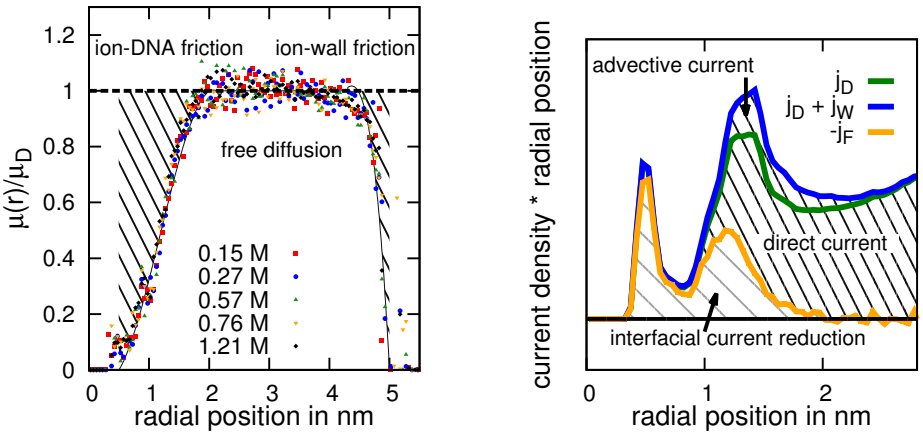
In the right panel of Fig. 23 we display the ion velocities $v^{\pm}$ of both ion species and the water velocity u as a function of the radial position r for the same data set. The continuum model (dotted curves) and the atomistic simulation data behave qualitatively similar. An electroosmotic flow of water is observed in the migration direction of the K^+ ions. Both ion species move significantly faster than the water whose velocity tends to zero in the atomistic model approximately at the inner boundary of the continuum model. The agreement of the hydrodynamic no-slip boundary between both models up to the size of a water molecule is remarkable.

Motivated by the simple decomposition of the ion velocity $v^{\pm} = \mu^{\pm}E + u$ in the continuum model, we define a position dependent mobility of the ions by computing their actual velocity minus the local water velocity, $\mu(r)^{\pm} = (v^{\pm} - u)/E$. In the region farthest away from both the pore wall and the DNA the mobility is constant, and we denote this constant value by μ_D, the free mobility. We find that its value is concentration dependent, and the trends of the experimental data in bulk electrolyte[69] are well reproduced (not shown). This was, however, not investigated in more detail. In the left panel of Fig. 24 we show the position dependent K^+ mobility for electrolyte concentrations between 0.15M and 1.2M, normalized by μ_D. All curves collapse onto a single master curve. The range of the mobility reduction is comparable with the effect found in Ref.[150]. The different geometries, however, make a direct comparison difficult.

Zhang and Shklovskii *e.g.*, introduced a friction coefficient between ions and DNA and[151], which they calculate based on hydrodynamic arguments, similar to those we

brought forward in Sec. 8.2. Comer *et al.* measured the diffusion coefficient of ions in the vicinity of the different nucleotides in atomistic simulations[150]. They found that the diffusion coefficient changes by more than a factor of two in the first nanometer from the nucleotide, independent of the specific base. We interpret the mobility reduction at the boundary as being caused by three factors. First, the major and minor groove produce a microscopically rough DNA surface, and ions trapped inside are virtually immobile. Second, the charge pattern on the DNA causes electrofriction[86,152] that can extend beyond the DNA backbone. Finally, hydrodynamic effects will slow down the motion of particles near a hydrodynamic boundary, as Comer *et al.* have shown.

In order to quantify how the mobility reduction next to the boundary affects the current, we split the current into three contributions: (1) the current I_D that would be expected if the mobility was constant and the water was immobile, (2) the convective current I_W due to the water flow, and (3) the (negative) current contribution I_F that is suppressed by interfacial friction. We express all of these in terms of current densities which we define as the direct current density $j_\mathrm{D} = c^+\mu_D^+E - c^-\mu_D^-E$, the convective current density (as in[153]) $j_\mathrm{W} = c^+u - c^-u$, and the interfacial current reduction density $j_\mathrm{F} = j - j_\mathrm{D} - j_\mathrm{W}$, where j is the total current density. The direct current depends, aside from the concentration dependence of the mobility, only on the number of ions of both species inside the pore. Therefore it contains both, the extra counterions that the DNA brings into the pore, and the ions expelled from the pore due to the DNA's finite volume.



(a) (b)

Figure 24: (a) Position dependent ion mobility normalized by the bulk mobility $(v-u)/\mu_D E$ for five different salt concentrations. Near the DNA and the wall a significant reduction is observed. (b) Observed current densities as a function of r. We distinguish the direct current j_D, the convective current j_W and the (negative) frictional current j_F. The direct current near the DNA is largely cancelled by friction.

In the right panel of Fig. 24 we report the direct and convective current density, and the density of interfacial current reduction. They are multiplied by the radius, so that the area under the graph corresponds to the magnitude of the associated current. The

area between the j_F curve (yellow) and the sum of j_D and j_W (blue) represents the total current. The shape of the direct current follows exactly the ion concentration reported in Fig. 23. The convective current $j_W = (c_+ - c_-)\, u$ is nonzero only in the regions with net charge, namely the Debye layer. It only occurs if a DNA is in the pore, and it is a positive modulation of the total current. In the proximity of the DNA, the conduction is largely suppressed. The fact that ions in the DNA's grooves are virtually immobile is reflected in j_D and j_F being almost equal. This however applies only to 1/3 of the counterions. Ions further away from the pore axis, also at $r > 1$ nm, have a considerably reduced mobility, resulting in a larger frictional current reduction.

We performed identical simulations also for a DNA-free pore. We investigate the relative current modulation $(I_{\text{with DNA}} - I_{\text{without DNA}})\,/I_{\text{without DNA}}$ as this quantity depends neither on the length of the investigated pore nor on the absolute value of the mobility of the ions. The total current was calculated from the displacement of the ions during a simulation run. In Fig. 25 the results are displayed together with the experimental data. The lines are only guides to the eye. The current obtained in the DNA-free simulations is interpolated by a third order polynomial to get a continuous function to determine the relative current modulation for every salt concentration. We decompose our data into a direct, an convective and a frictional component. For all salt concentrations the current modulation

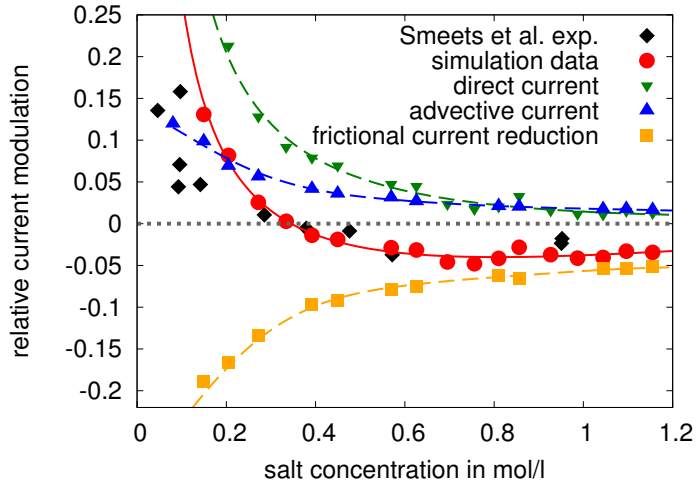


Figure 25: Relative current modulation caused by the DNA as a function of salt concentration for the experiments (diamonds) and simulations (circles). The quantitative agreement between simulations and experiments is good, especially the location of the crossover point between enhancement and reduction. We decomposed the total current into direct current (downward triangles), convective current (upward triangles), and friction (squares) and consider their contribution to the modulation independently. The current reduction due to friction is responsible for observing an overall current reduction at large salt concentrations.

from the simulations are in good agreement with the experimental data. For high salt conditions a current reduction of around 5% is observed. The experiments suggest a somewhat lower reduction, however with uncertainties similar to the deviations between experiments and simulation. For low salt concentrations the predicted current enhancement is larger

than experimentally observed. This can be justified by noting that the experimentally used nanopores are not very long (around 40 nm) and have an hourglass shape. Only the radius at the constriction corresponds to the 10 nm used here. When the pore is wider the relative enhancement or blockade can be expected to be smaller. The crossover point between enhancement and blockade occurs at an electrolyte concentration of 0.3 M, agreeing with the experiments.

The observed direct current is always larger in presence of the DNA since the number of ions in the pore at equal reservoir concentration is always larger with DNA in the investigated concentration range up to 1.2 M. The continuum model predicts this to be the case up to 0.85 M. Thus the obstruction of the cross section due to the DNA's finite volume is always overcompensated by the presence of the Debye layer. The convective contribution to the current only appears in the presence of a DNA strand and is always positive. The effect that is responsible for the crossover effect, however, can be clearly identified. Without the mobility reduction at the surface of the DNA no current blockades would be observed.

This comparison of LB simulations to atomistic simulations, experimental results and numerical calculations demonstrate both the strengths and drawbacks associated with our approach. The Lattice-Boltzmann simulations did an excellent job at reproducing the bulk behaviour of the system. The obvious drawback of the LB simulations was their failure to reproduce the increased hydrodynamic drag experienced by ions near the surface. While this shows the need to tune the LB simulations in order to reproduce the correct physical behaviour, they still allow for a dramatic increase in system size, something which is not possible with the atomistic approach. In addition it gives us important insight in how to improve the continuum model. More details about the results presented in this chapter can be found in the original publications[154,147,66]

9 Conclusion and Outlook

In this book chapter we have looked at the use of molecular dynamics simulations combined with a Lattice-Boltzmann fluid for the simulation of electrokinetic phenomena in soft matter. The section 2 started with an introduction to the basic theoretical framework which is used to understand such phenomena, followed by a detailed explanation of the simulation method in section 4. Next we looked at how the MD-LB hybrid simulations performed on a number of simple systems.

In the first system we examined in section 5, the electrophoresis of simple colloids in bulk solution, the method was able to accurately reproduce experimental colloid mobilities. While this method has proved successful for relatively small colloids from one to ten nanometers in size, the raspberry model used will probably struggle with increasing colloidal size. An interesting future avenue of investigation would be to model the colloid using the boundary conditions suggested by Ladd *et al.* [155], allowing for the simulation of significantly larger colloids. The next system investigated in section 6 was the electrophoresis of a single polyelectrolyte in bulk solution, where we again saw excellent agreement with experimental results. One possibility for new simulations in this area would be to simulate a polyelectrolyte under conditions more typical in experiment, such as in a polymer network or microfluidic device. The next example was the electrophoresis of a polyelectrolyte-grafted colloid in section 7. The simulation results showed excellent agreement with numerical calculations based on the standard electrokinetic model.

Furthermore, the system displayed a mobility reversal with increasing concentration of *monovalent* salt. Here, the large number of parameters in the system leaves a large part of the phase space yet to be explored. There are also a wealth of experimental results which could be compared to. The final system which we examined in section 8 was the role of the presence of a charged dsDNA on the conductivity of a nanopore. In this example we initially observed a significant deviation from the experimental results due to the neglect of the change of the ion friction near the DNA, which has been detected using all-atom simulations. An incorporation of this back into the LB scheme is possible and leads to a perfect agreement with the all-atom simulations[156]. This opens up the possibility to look at finite length pores and the role of the DNA's motion on the conductivity using the LB scheme, since those aspects are out of range for an all-atom treatment.

Overall, the simulations presented in this contribution conclusively demonstrate the ability of Lattice-Boltzmann to produce meaningful results for soft matter systems. In the near future, we foresee that Lattice-Boltzmann simulations will be able to accurately model an ever increasing number of soft matter systems as processor speeds and the amount of memory available continue to increase. This will no doubt be coupled with a transition from a time where simulations are used to understand existing experimental results, to an era where simulations are used to help design optimized experimental setups.

Acknowledgments

We are very grateful to Vladimir Lobaskin, Burkhard Dünweg, Kai Grass, Shervin Raafatnia and Stefan Kesselheim from whose articles and theses much of the contents of this book chapter was copied almost verbatim. This work is in fact a compendium of their important contributions to the development of Lattice-Boltzmann simulations in the context of the electrokinetics of soft matter systems. We furthermore acknowledge funding from the DFG through the SFB 716 and the project HO 1108/22-1.

References

1. P. Ahlrichs and B. Dünweg, *Simulation of a single polymer chain in solution by combining lattice Boltzmann and molecular dynamics*, Journal of Chemical Physics, **111**, no. 17, 8225–8239, 1999.
2. J. Horbach and S. Succi, *Lattice Boltzmann versus molecular dynamics simulation of nanoscale hydrodynamic flows*, Physical Review Letters, **96**, no. 22, 224503, 2006.
3. D. Röhm and A. Arnold, *Lattice Boltzmann simulations on GPUs with ESPResSo*, The European Physical Journal Special Topics, **210**, 89–100, 2012.
4. K. C. Grass, *Towards realistic modelling of free solution electrophoresis: a case study on charged macromolecules*, PhD thesis, Goethe-Universität Frankfurt am Main, 2008.
5. LD Landau and EM Lifshitz, *Statistische Physik, vol. 5*, Akademie-Verlag, Berlin, 1987.
6. Shervin Raafatnia, *The Electrophoretic Mobility of Bare and Soft Spherical Colloids: a Molecular Dynamics Study*, PhD thesis, Institute for Computational Physics, University of Stuttgart, Germany, 2015.

7. Á V. Delgado, F. González Caballero, R. J. Hunter, L. K. Koopal, and J. Lyklema, *Measurement and interpretation of electrokinetic phenomena - (IUPAC technical report)*, 2005.
8. B. Dünweg, V. Lobaskin, K. Seethalakshmy-Hariharan, and C. Holm, *Colloidal electrophoresis: scaling analyisis, Green-Kubo relation, and numerical results*, Journal of Physics: Condensed Matter, **20**, 404214, 2008.
9. W. B. Russel, D. A. Saville, and W. R. Schowalter, *Colloidal Dispersions*, Cambridge University Press, Cambridge, UK, 1989.
10. F.A Morrison, Jr., *Electrophoresis of a particle of arbitrary shape*, Journal of Colloid and Interface Science, **34**, no. 2, 210–214, 1970.
11. Richard Wyndham O'Brien and Robert John Hunter, *The electrophoretic mobility of large colloidal particles*, Canadian Journal of Chemistry, **59**, no. 13, 1878–1887, July 1981.
12. Mike P. Allen and Dominik J. Tildesley, *Computer Simulation of Liquids*, Oxford Science Publications. Clarendon Press, Oxford, 1 edition, 1987.
13. D. Frenkel, *Colloidal systems: Playing Tricks with Designer "Atoms"*, Science, **296**, 65, Nov. 2002.
14. J. D. Weeks, D. Chandler, and H. C. Andersen, *Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids*, Journal of Chemical Physics, **54**, 5237, 1971.
15. T. Soddemann, B. Dünweg, and K. Kremer, *A generic computer model for amphiphilic systems*, European Physical Journal E: Soft Matter, **6**, 409, 2001.
16. A. Arnold and C. Holm, "Efficient methods to compute long range interactions for soft matter systems", in: Advanced Computer Simulation Approaches for Soft Matter Sciences II, C. Holm and K. Kremer, (Eds.), vol. II of *Advances in Polymer Sciences*, pp. 59–109. Springer, Berlin, 2005.
17. S. W. de Leeuw, J. W. Perram, and E. R. Smith, *Simulation of Electrostatic Systems in Periodic Boundary Conditions. I. Lattice Sums and Dielectric Constants*, Proceedings of the Royal Society of London. Series A: Mathematics and Physical Sciences, **373**, no. 1752, 27–56, Oct. 1980.
18. P. P. Ewald, *Die Berechnung optischer und elektrostatischer Gitterpotentiale*, Annales de Physique, **369**, no. 3, 253–287, 1921.
19. J. W. Eastwood, R. W. Hockney, and D. N. Lawrence, *P3M3DP–The three-dimensional periodic particle-particle/ particle-mesh program*, Computer Physics Communications, **19**, no. 2, 215–261, 1980.
20. R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles*, IOP, London, 1988.
21. T. Darden, D. York, and L. Pedersen, *Particle Mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems*, Journal of Chemical Physics, **98**, 10089–10092, 1993.
22. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. Pedersen, *A smooth Particle Mesh Ewald method*, Journal of Chemical Physics, **103**, 8577, 1995.
23. M. Deserno and C. Holm, *How to mesh up Ewald sums. I. A theoretical and numerical comparison of various particle mesh routines*, Journal of Chemical Physics, **109**, 7678, 1998.
24. L. Greengard and V. Rhoklin, *A fast algorithm for Particle Simulations*, Journal of Computational Physics, **73**, 325, 1987.

25. Frederick S. Lee and Arieh Warshel, *A local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations*, Journal of Chemical Physics, **97**, no. 5, 3100–3107, 1992.
26. Axel Arnold, Florian Fahrenberger, Christian Holm, Olaf Lenz, Matthias Bolten, Holger Dachsel, Rene Halver, Ivo Kabadshow, Franz Gähler, Frederik Heber, Julian Iseringhausen, Michael Hofmann, Michael Pippig, Daniel Potts, and Godehard Sutmann, *Comparison of scalable fast methods for long-range interactions*, Phys. Rev. E, **88**, 063308, Dec 2013.
27. M. Deserno and C. Holm, *How to mesh up Ewald sums. II. An accurate error estimate for the Particle-Particle-Particle-Mesh algorithm*, Journal of Chemical Physics, **109**, 7694, 1998.
28. V. Ballenegger, J. J. Cerdà, O. Lenz, and Ch. Holm, *The optimal P3M algorithm for computing electrostatic energies in periodic systems*, Journal of Chemical Physics, **128**, no. 3, 034109, 2008.
29. V. Ballenegger, J. J. Cerdà, and C. Holm, *How to Convert SPME to P3M: Influence Functions and Error Estimates*, Journal of Chemical Theory and Computation, **8**, no. 3, 936–947, 2012.
30. S. Tyagi, A. Arnold, and C. Holm, *ICMMM2D: An accurate method to include planar dielectric interfaces via image charge summation*, Journal of Chemical Physics, **127**, 154723, 2007.
31. Sandeep Tyagi, Axel Arnold, and Christian Holm, *Electrostatic layer correction with image charges: A linear scaling method to treat slab 2D + h systems with dielectric interfaces*, Journal of Chemical Physics, **129**, no. 20, 204102, 2008.
32. C. Tyagi, M. Süzen, M. Sega, M. Barbosa, S. Kantorovich, and C. Holm, *An iterative, fast, linear-scaling method for computing induced charges on arbitrary dielectric boundaries*, Journal of Chemical Physics, **132**, 1154112, 2010.
33. Axel Arnold, Konrad Breitsprecher, Florian Fahrenberger, Stefan Kesselheim, Olaf Lenz, and Christian Holm, *Efficient Algorithms for Electrostatic Interactions Including Dielectric Contrasts*, Entropy, **15**, no. 11, 4569–4588, 2013.
34. Florian Fahrenberger, Zhenli Xu, and Christian Holm, *Simulation of electric double layers around charged colloids in aqueous solution of variable permittivity*, J. Chem. Phys., **141**, no. 6, 2014.
35. F. Fahrenberger and C. Holm, *Computing the Coulomb interaction in inhomogeneous dielectric media via a local electrostatics lattice algorithm*, Phys. Rev. E, **90**, 063304, Dec 2014.
36. A. Malevanets and R. Kapral, *Mesoscopic model for solvent dynamics*, Journal of Chemical Physics, **110**, 8605–8613, May 1999.
37. T. Ihle and D. M. Kroll, *Stochastic rotation dynamics. I. Formalism, Galilean invariance, and Green-Kubo relations*, Phys. Rev. E, **67**, no. 6, 066705, June 2003.
38. T. Ihle and D. M. Kroll, *Stochastic rotation dynamics. II. Transport coefficients, numerics, and long-time tails*, Phys. Rev. E, **67**, no. 6, 066706, June 2003.
39. Pep Español and Patrick Warren, *Statistical mechanics of Dissipative Particle Dynamics*, Europhysics Letters, **30**, 191, 1995.
40. P. Español, *Fluid particle dynamics: A synthesis of dissipative particle dynamics and smoothed particle dynamics*, EPL (Europhysics Letters), **39**, no. 6, 605, 1997.
41. P. J. Hoogerbrugge and J. M. V. A. Koelman, *Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics*, Europhysics Letters, **19**, no. 3, 155–160, 1992.

42. G. R. McNamara and G. Zanetti, *Use of the Boltzmann-equation to simulate lattice-gas automata*, Physical Review Letters, **61**, no. 20, 2332–2335, Nov.14 1988.
43. R. Benzi, S. Succi, and M. Vergassola, *The lattice Boltzmann equation: theory and applications*, Physics Reports, **222**, no. 3, 145–197, 1992.
44. A. J. C. Ladd, *Numerical simulations of particulate suspensions via a discretized Boltzmann equation. Part 1. Theoretical foundation*, Journal of Fluid Mechanics, **271**, 285–309, 1994.
45. S. Chen and G. D. Doolen, *Lattice Boltzmann method for fluid flows*, Annual Review of Fluid Mechanics, **30**, no. 1, 329–364, 1998.
46. J. M. Yeomans, *Mesoscale simulations: Lattice Boltzmann and particle algorithms*, Physica A, **369**, 159–184, 2006.
47. G. Gompper and T. Ihle, *Advanced computer simulation approaches for soft matter sciences I*, chapter Multi-particle collision dynamics: A particle-based mesoscale simulation approach to the hydrodynamics of complex fluids, p. 6, Springer, 2009.
48. B. Dünweg and A. J. C. Ladd, "Lattice boltzmann simulations of soft matter systems", in: Advanced Computer Simulation Approaches for Soft Matter Sciences III, vol. 221 of *Advances in Polymer Science*, pp. 89–166. Springer-Verlag Berlin, Berlin, Germany, 2009.
49. Christoph Junghans, Matej Praprotnik, and Kurt Kremer, *Transport properties controlled by a thermostat: An extended dissipative particle dynamics thermostat*, Soft Matter, **4**, no. 1, 156–161, 2008.
50. C. M. Pooley and J. M. Yeomans, *Kinetic Theory Derivation of the Transport Coefficients of Stochastic Rotation Dynamics*, The Journal of Physical Chemistry B, **109**, no. 14, 6505–6513, 2005.
51. J. Smiatek, M. P. Allen, and F. Schmid, *Tunable-slip boundaries for coarse-grained simulations of fluid flow*, European Physical Journal E: Soft Matter, **26**, 115–122, 2008.
52. AJC Ladd and R. Verberg, *Lattice-Boltzmann simulations of particle-fluid suspensions*, J. Stat. Phys., **104**, no. 5, 1191–1251, 2001.
53. Jens Smiatek, Marcello Sega, Christian Holm, Ulf D. Schiller, and Friederike Schmid, *Mesoscopic simulations of the counterion-induced electro-osmotic flow: A comparative study*, Journal of Chemical Physics, **130**, no. 244702, 06 2009.
54. Jens Smiatek and Friederike Schmid, *Polyelectrolyte Electrophoresis in Nanochannels: A Dissipative Particle Dynamics Simulation*, Journal of Physical Chemistry B, **114**, no. 19, 6266–6272, 2010.
55. Jens Smiatek and Friederike Schmid, *Mesoscopic simulations of electroosmotic flow and electrophoresis in nanochannels*, Comp. Phys. Comm., **182**, no. 9, 1941–1944, 2011, Computer Physics Communications Special Edition for Conference on Computational Physics Trondheim, Norway, June 23-26, 2010.
56. S. Frank and R. G. Winkler, *Polyelectrolyte electrophoresis: field effects and hydrodynamic interactions*, Europhysics Letters, **83**, 38004, 2008.
57. Sandra Frank and Roland G. Winkler, *Mesoscale hydrodynamic simulation of short polyelectrolytes in electric fields*, Journal of Chemical Physics, **131**, no. 23, Dec. 21 2009.
58. Fabrizio Capuani, Ignacio Pagonabarraga, and Daan Frenkel, *Discrete solution of the electrokinetic equations*, The Journal of chemical physics, **121**, 973, 2004.

59. I. Pagonabarraga, B. Rotenberg, and D. Frenkel, *Recent advances in the modelling and simulation of electrokinetic effects: bridging the gap between atomistic and macroscopic descriptions*, Physical Chemistry Chemical Physics, **12**, 9566–9580, 2010.
60. Giovanni Giupponi and Ignacio Pagonabarraga, *Colloid Electrophoresis for Strong and Weak Ion Diffusivity*, Physical Review Letters, **106**, no. 24, 248304–, June 2011.
61. P. L. Bhatnagar, E. P. Gross, and M. Krook, *A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems*, Physical Review, **94**, no. 3, 511, 1954.
62. U. D. Schiller, *Thermal fluctuations and boundary conditions in the lattice Boltzmann method*, PhD thesis, Johannes Gutenberg-Universitaet Mainz, Fachbereich 08: Physik, Mathematik und Informatik, 2008.
63. D. d'Humieres, *Multiple–relaxation–time lattice Boltzmann models in three dimensions*, Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, **360**, no. 1792, 437–451, 2002.
64. R. Adhikari, K. Stratford, ME Cates, and AJ Wagner, *Fluctuating Lattice Boltzmann*, Europhysics Letters, **71**, 473, 2005.
65. B. Dünweg, U. Schiller, and A. J. C. Ladd, *Statistical mechanics of the fluctuating Lattice-Boltzmann equation*, Phys. Rev. E, **76**, 36704, 2007.
66. Stefan Kesselheim, *Simulations of DNA Translocation through nanopores*, PhD thesis, Universitaet Stuttgart, 2014.
67. J. T. Padding and A. A. Louis, *Hydrodynamic interactions and Brownian forces in colloidal suspensions: Coarse-graining over time and length scales*, Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, **74**, no. 3, 031402, 2006.
68. Lars. Onsager, *Theories of Concentrated Electrolytes.*, Chemical Reviews, **13**, no. 1, 73–89, 1933.
69. A. L. Horvath, *Handbook of Aqueous Electrolyte Solutions*, Ellis Horwood Ltd., Chichester, 1 edition, 1985.
70. P. G. Righetti, (Ed.), *Capillary Electrophoresis in Analytical Biotechnology*, CRC Press, Boca Raton, 1996.
71. H. Cottet and P. Gareil, *From small charged molecules to oligomers: a semiempirical approach to the modeling of actual mobility in free solution.*, Electrophoresis, **21**, no. 8, 1493–1504, May 2000.
72. H. Cottet, P. Gareil, O. Theodoly, and C. E. Williams, *A semi-empirical approach to the modeling of the electrophoretic mobility in free solution: Application to polystyrenesulfonates of various sulfonation rates*, Electrophoresis, **21**, 3529–3540, 2000.
73. U. Böhme and U. Scheler, *Effective charge of poly(styrenesulfonate) and ionic strength - an electrophoresis NMR investigation*, Colloids and Surfaces A, **222**, no. 1-3, 35–40, July 2003.
74. N. C. Stellwagen, C. Gelfi, and P. G. Righetti, *The free solution mobility of DNA*, Biopolymers, **42**, no. 6, 687–703, Nov. 1997.
75. N. C. Stellwagen, S. Magnusdottir, C. Gelfi, and P. G. Righetti, *Measuring the translational diffusion coefficients of small DNA molecules by capillary electrophoresis*, Biopolymers, **58**, no. 4, 390–397, Apr. 2001.
76. R. R. Netz, *Nonequilibrium unfolding of polyelectrolyte condensates in electric fields*, Physical Review Letters, **90**, no. 12, 128104, Mar. 2003.

77. R. R. Netz, *Polyelectrolytes in electric fields*, Journal of Physical Chemistry B, **107**, no. 32, 8208–8217, Aug. 2003.
78. V. Lobaskin, B. Dünweg, M. Medebach, T. Palberg, and C Holm, *Electrophoresis of Colloidal Dispersions in the Low-Salt Regime*, Physical Review Letters, **98**, 176105, Apr. 2007.
79. V. Lobaskin, B. Dünweg, and C. Holm, *Electrophoretic mobility of a charged colloidal particle: A computer simulation study*, Journal of Physics: Condensed Matter, **16**, no. 38, S4063–S4073, Sept. 2004.
80. B Dünweg, V Lobaskin, K Seethalakshmy-Hariharan, and C Holm, *Colloidal electrophoresis: scaling analysis, Green-Kubo relation, and numerical results*, Journal of Physics: Condensed Matter, **20**, no. 40, 404214, 2008.
81. V. Lobaskin and B. Dünweg, *A new model for simulating colloidal dynamics*, New Journal of Physics, **6**, 54, May 2004.
82. Joost de Graaf and et al., work submitted, 2015.
83. Frédéric Tessier and Gary W. Slater, *Effective Debye length in closed nanoscopic systems: A competition between two length scales*, Electrophoresis, **27**, no. 3, 686–693, 2006.
84. P. H. Wiersema, A. L. Loeb, and J. Th.G. Overbeek, *Calculation of Electrophoretic Mobility of a spherical Colloid Particle*, Journal of Colloid and Interface Science, **22**, 78, 1966.
85. R. W. O'Brien and L. R. White, *Electrophoretic Mobility of a spherical colloidal Particle*, J. Chem. Soc. Faraday Trans., **74**, no. 2, 1607, 1978.
86. RR Netz, *Electrofriction and dynamic stern layers at planar charged surfaces*, Physical Review Letters, **91**, no. 13, 138101, 2003.
87. K Grass and C. Holm, *Mesoscale modelling of polyelectrolyte electrophoresis*, Faraday Discussions, **144**, 57–70, 2010.
88. J. L. Viovy, *Electrophoresis of DNA and other polyelectrolytes: Physical mechanisms*, Reviews of Modern Physics, **72**, no. 3, 813–872, 2000.
89. G. W. Slater, C. Holm, M. V. Chubynsky, H. W. de Haan, A. Dubé, K. Grass, O. A. Hickey, C. Kingsburry, D. Sean, T. N. Shendruk, and L. Zhan, *Modeling the separation of macromolecules: A review of current computer simulation methods*, Electrophoresis, **30**, no. 5, 792–818, Mar. 2009.
90. K. Grass, U. Böhme, U. Scheler, H. Cottet, and C. Holm, *Importance of Hydrodynamic Shielding for the Dynamic Behavior of Short Polyelectrolyte Chains*, Physical Review Letters, **100**, 096104, 2008.
91. K. Grass and C. Holm, *On the importance of hydrodynamic interactions in polyelectrolyte electrophoresis*, Journal of Physics: Condensed Matter, **20**, 494217, 2008.
92. D. A. Hoagland, E. Arvanitidou, and C. Welch, *Capillary electrophoresis measurements of the free solution mobility for several model polyelectrolyte systems*, Macromolecules, **32**, no. 19, 6180–6190, Sept. 1999.
93. G. Manning, *Limiting Laws and Counterion Condensation in Polyelectrolyte Solutions I. Colligative Properties*, Journal of Chemical Physics, **51**, 924–933, 1969.
94. F. Oosawa, *Polyelectrolytes*, Marcel Dekker, New York, 1971.
95. Shervin Raafatnia, Owen A. Hickey, and Christian Holm, *Mobility Reversal of Polyelectrolyte-Grafted Colloids in Monovalent Salt Solutions*, Phys. Rev. Lett., **113**, 238301, Dec 2014.

96. M. Elimelech and C. R. O'Melia, *Effect of electrolyte type on the electrophoretic mobility of polystyrene latex colloids*, Colloids and Surfaces, **44**, 165–178, 1990.
97. M. Quesada-Perez, A. Martin-Molina, F. Galisteo-Gonzalez, and R. Hidalgo-Alvarez, *Electrophoretic mobility of model colloids and overcharging: theory and experiment*, Molecular Physics, **100**, no. 18, 3029–3039, 2002.
98. Christophe Labbez, Andrá Nonat, Isabelle Pochard, and Bo Jonsson, *Experimental and theoretical evidence of overcharging of calcium silicate hydrate*, Journal of Colloid and Interface Science, **309**, no. 2, 303–307, 2007.
99. Alberto Martin-Molina, Jose Alberto Maroto-Centeno, Roque Hidalgo-Alvarez, and Manuel Quesada-Perez, *Charge reversal in real colloids: Experiments, theory and simulations*, Colloids Surf. A, **319**, no. 1-3, 103–108, 2008.
100. Christian Schneider, Mathias Hanisch, Bastian Wedel, Arben Jusufi, and Matthias Ballauff, *Experimental study of electrostatically stabilized colloidal particles: Colloidal stability and charge reversal*, Journal of Colloid and Interface Science, **358**, no. 1, 62–67, June 2011.
101. Motohiko Tanaka and Alexander Yu Grosberg, *Electrophoresis of charge inverted macroion complex: Molecular dynamics study*, Euro. Phys. J. E, **7**, no. 4, 371–379, 2002.
102. Pai-Yi Hsiao and Erik Luijten, *Salt-Induced Collapse and Reexpansion of Highly Charged Flexible Polyelectrolytes*, Physical Review Letters, **97**, 148301, 2006.
103. A. Diehl and Y. Levin, *Smoluchowski equation and the colloidal charge reversal*, Journal of Chemical Physics, **125**, no. 5, 054902, Aug. 2006.
104. Anna Kubíčková, Tomáš Křížek, Pavel Coufal, Mario Vazdar, Erik Wernersson, Jan Heyda, and Pavel Jungwirth, *Overcharging in Biological Systems: Reversal of Electrophoretic Mobility of Aqueous Polyaspartate by Multivalent Cations*, Physical Review Letters, **108**, 186101, May 2012.
105. Ilya Semenov, Shervin Raafatnia, Marcello Sega, Vladimir Lobaskin, Christian Holm, and Friedrich Kremer, *Electrophoretic mobility and charge inversion of a colloidal particle studied by single-colloid electrophoresis and molecular dynamics simulations*, Phys. Rev. E, **87**, 022302, Feb 2013.
106. Shervin Raafatnia, Owen A. Hickey, Marcello Sega, and Christian Holm, *Computing the Electrophoretic Mobility of Large Spherical Colloids by Combining Explicit Ion Simulations with the Standard Electrokinetic Model*, Langmuir, **30**, no. 7, 1758–1767, 2014.
107. Hiroyuki Ohshima, *Dynamic Electrophoretic Mobility of a Soft Particle*, Journal of Colloid and Interface Science, **233**, no. 1, 142 – 152, 2001.
108. R. J. Hill and D. A. Saville, *'Exact' solutions of the full electrokinetic model for soft spherical colloids: Electrophoretic mobility*, Colloids Surf. A, **267**, no. 1-3, 31–49, 2005.
109. Hiroyuki Ohshima, *Electrokinetics of soft particles*, **285**, no. 13, 1411–1421–, 2007.
110. Nikolay Houbenov, Sergiy Minko, and Manfred Stamm, *Mixed polyelectrolyte brush from oppositely charged polymers for switching of surface charge and composition in aqueous environment*, Macromolecules, **36**, no. 16, 5897–5901, 2003.
111. L.A Rosen and D.A Saville, *The electrokinetic response of surface-modified polymer latexes: Effects of grafted water-soluble polymer and heat treatment*, Journal of Colloid and Interface Science, **149**, no. 2, 542 – 552, 1992.

112. Hiroshi Hayashi, Satoshi Tsuneda, Akira Hirata, and Hiroshi Sasaki, *Soft particle analysis of bacterial cells and its interpretation of cell adhesion behaviors in terms of DLVO theory*, Colloids and Surfaces B: Biointerfaces, **22**, no. 2, 149 – 157, 2001.
113. Alexis J. de Kerchove and Menachem Elimelech, *Relevance of Electrokinetic Theory for "Soft" Particles to Bacterial Cells: Implications for Bacterial Adhesion*, Langmuir, **21**, no. 14, 6462–6472, 2005, PMID: 15982054.
114. Kimiko Makino and Hiroyuki Ohshima, *Soft particle analysis of electrokinetics of biological cells and their model systems*, Science and Technology of Advanced Materials, **12**, no. 2, 023001, 2011.
115. Jérôme F.L. Duval and Fabien Gaboriaud, *Progress in electrohydrodynamics of soft microbial particle interphases*, Current Opinion in Colloid & Interface Science, **15**, no. 3, 184–195, June 2010.
116. S. Levine, M. Levine, K.A. Sharp, and D.E. Brooks, *Theory of the electrokinetic behavior of human erythrocytes*, Biophysical Journal, **42**, 127 – 135, 1983.
117. Rolf Bos, Henny C. van der Mei, and Henk J. Busscher, *'Soft-particle' analysis of the electrophoretic mobility of a fibrillated and non-fibrillated oral streptococcal strain: Streptococcus salivarius*, Biophysical Chemistry, **74**, no. 3, 251 – 255, 1998.
118. Hans J. Kaper, Henk J. Busscher, and Willem Norde, *Characterization of poly(ethylene oxide) brushes on glass surfaces and adhesion of Staphylococcus epidermidis*, Journal of Biomaterials Science, Polymer Edition, **14**, no. 4, 313–324, 2003.
119. Grégoire Danger, Michel Ramonda, and Hervé Cottet, *Control of the EOF in CE using polyelectrolytes of different charge densities*, Electrophoresis, **28**, no. 6, 925–931, 2007.
120. Owen A. Hickey, Christian Holm, James L. Harden, and Gary W. Slater, *Influence of Charged Polymer Coatings on Electro-Osmotic Flow: Molecular Dynamics Simulations*, Macromolecules, **44**, 9455–9463, 2011.
121. Il Sang Lee, Min Seong Cho, and Hyoung Jin Choi, *Preparation of polyaniline coated poly(methyl methacrylate) microsphere by graft polymerization and its electrorheology*, Polymer, **46**, no. 4, 1317 – 1321, 2005, Collected papers from PDM 2004. Polymers in Dispersed Media - Colloids: from Preparation to Application.
122. Young Dae Kim and Dong Hyup Park, *Transient overshoot of the electrorheological responses of conducting polymer-coated polyethylene suspensions in mineral oil*, Synthetic Metals, **142**, no. 1-3, 147 – 151, 2004.
123. Vladimir P. Torchilin, *Targeted pharmaceutical nanocarriers for cancer therapy and imaging*, The AAPS Journal, **9**, no. 2, E128–E147, 2007.
124. Khin Yin Win and Si-Shen Feng, *Effects of particle size and surface coating on cellular uptake of polymeric nanoparticles for oral delivery of anticancer drugs*, Biomaterials, **26**, no. 15, 2713 – 2722, 2005.
125. Veska Toncheva, Margreet A Wolfert, Philip R Dash, David Oupicky, Karel Ulbrich, Leonard W Seymour, and Etienne H Schacht, *Novel vectors for gene delivery formed by self-assembly of DNA with poly(l-lysine) grafted with hydrophilic polymers*, Biochimica et Biophysica Acta (BBA) - General Subjects, **1380**, no. 3, 354 – 368, 1998.
126. H Ohshima, M Nakamura, and T Kondo, *Electrophoretic mobility of colloidal particles coated with a layer of adsorbed polymers*, Colloid and Polymer Science, **270**, no. 9, 873–877, 1992.

127. Hiroyuki Ohshima, *Electrokinetic phenomena of soft particles*, Current Opinion in Colloid & Interface Science, **18**, no. 2, 73 – 82, 2013.
128. J.J. López-García, C. Grosse, and J. Horno, *Numerical study of colloidal suspensions of soft spherical particles using the network method: 1. DC electrophoretic mobility*, Journal of Colloid and Interface Science, **265**, no. 2, 327–340, Sept. 2003.
129. Stanislav S. Dukhin, Ralf Zimmermann, Jérôme F. Duval, and Carsten Werner, *On the applicability of the Brinkman equation in soft surface electrokinetics*, Journal of Colloid and Interface Science, **350**, no. 1, 1–4, Oct. 2010.
130. Alexander C. Barbati and Brian J. Kirby, *Soft diffuse interfaces in electrokinetics - theory and experiment for transport in charged diffuse layers*, Soft Matter, **8**, no. 41, 10598–10613, 2012.
131. Jérôme F. L. Duval and Hiroyuki Ohshima, *Electrophoresis of Diffuse Soft Particles*, Langmuir, **22**, no. 8, 3533–3546, march 2006.
132. D.A. Saville, *Electrokinetic Properties of Fuzzy Colloidal Particles*, Journal of Colloid and Interface Science, **222**, no. 1, 137 – 145, 2000.
133. H. J. Limbach, A. Arnold, B. A. Mann, and C. Holm, *ESPResSo – An Extensible Simulation Package for Research on Soft Matter Systems*, Comp. Phys. Comm., **174**, no. 9, 704–727, May 2006.
134. A. Arnold, O. Lenz, S. Kesselheim, R. Weeber, F. Fahrenberger, D. Röhm, P. Košovan, and C. Holm, *ESPResSo 3.1 — Molecular Dynamics Software for Coarse-Grained Models*, in: Meshfree Methods for Partial Differential Equations VI, M. Griebel and M. A. Schweitzer, (Eds.), vol. 89 of *Lecture Notes in Computational Science and Engineering*, pp. 1–23, Springer. 2013.
135. RJ Hill, DA Saville, and WB Russel, *Electrophoresis of spherical polymer-coated colloidal particles*, Journal of Colloid and Interface Science, **258**, no. 1, 56–74, Feb. 1 2003.
136. Hiroyuki Ohshima, *Electrophoretic Mobility of a Spherical Colloidal Particle in a Salt-Free Medium*, Journal of Colloid and Interface Science, **248**, no. 2, 499–503, 2002.
137. Stanislav S. Dukhin, Ralf Zimmermann, and Carsten Werner, *Electrophoresis of soft particles at high electrolyte concentrations: An interpretation by the Henry theory*, Journal of Colloid and Interface Science, **313**, no. 2, 676 – 679, 2007.
138. Hiroyuki Ohshima, *Electrophoretic mobility of soft particles*, Colloids and Surfaces A: Physicochemical and Engineering Aspects, **103**, no. 3, 249 – 255, 1995.
139. J. L. Harden, D. Long, and A. Ajdari, *Influence of End-Grafted Polyelectrolytes on Electro-Osmosis along Charged Surfaces*, Langmuir, **17**, no. 3, 705–715, Jan. 2001.
140. M. Deserno, C. Holm, J. Blaul, M. Ballauff, and M. Rehahn, *The Osmotic Coefficient of Rod-like Polyelectrolytes: Computer Simulation, Analytical Theory, and Experiment*, European Physical Journal E: Soft Matter, **5**, 97–103, 2001.
141. G.S. Perkins and R.B. Jones, *Hydrodynamic interaction of a spherical particle with a planar boundary: II. Hard wall*, Physica A: Statistical Mechanics and its Applications, **189**, 447 – 477, 1992.
142. David Van Der Spoel, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E. Mark, and Herman J. C. Berendsen, *GROMACS: Fast, flexible, and free*, Journal of Computational Chemistry, **26**, no. 16, 1701–1718, 2005.
143. Ralph M. M. Smeets, Ulrich F. Keyser, Diego Krapf, Meng-Yue Wu, Nynke H. Dekker, and Cees Dekker, *Salt dependence of ion transport and DNA translocation through solid-state nanopores*, Nano Letters, **6**, 89–95, 2006.

144. B. Luan and A. Aksimentiev, *Electro-osmotic screening of the DNA charge in a nanopore*, Physical Review E, **78**, no. 2, 021912, 2008.
145. B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*, Journal of Chemical Theory and Computation, **4**, no. 3, 435–447, 2008.
146. H. J. C. Berendsen, D. van der Spoel, and R. van Drunen, *GROMACS: A message-passing parallel molecular dynamics implementation*, Computer Physics Communications, **91**, no. 1-3, 43–56, Sept. 1995.
147. Stefan Kesselheim, Wojciech Müller, and Christian Holm, *Origin of Current Blockades in Nanopore Translocation Experiments*, Physical Review Letters, **112**, 018101, Jan 2014.
148. Michael Feig and B. Montgomery Pettitt, *Sodium and chlorine ions as part of the DNA solvation shell*, Biophysical Journal, **77**, 1769–1781, 1999.
149. Christopher Maffeo, Robert Schöpflin, Hergen Brutzer, René Stehr, Aleksei Aksimentiev, Gero Wedemann, and Ralf Seidel, *DNA–DNA interactions in tight supercoils are described by a small effective charge density*, Physical Review Letters, **105**, no. 15, 158101, 2010.
150. Jeffrey Comer and Aleksei Aksimentiev, *Predicting the DNA sequence dependence of nanopore ion current using atomic-resolution Brownian dynamics*, The Journal of Physical Chemistry C, **116**, no. 5, 3376–3393, 2012.
151. J. Zhang and B.I. Shklovskii, *Effective charge and free energy of DNA inside an ion channel*, Phys. Rev. E, **75**, no. 2, 21906, 2007.
152. Yong Woon Kim and Roland R Netz, *Electro-osmosis at inhomogeneous charged surfaces: Hydrodynamic versus electric friction*, The Journal of chemical physics, **124**, 114709, 2006.
153. Payam Rowghanian and Alexander Y. Grosberg, *Two cases of reciprocal relations for electric and hydrodynamic currents: A rigid polymer in a nano-channel and a polyelectrolyte gel*, The Journal of Chemical Physics, **139**, no. 2, –, 2013.
154. Stefan Kesselheim and Christian Holm, *Modeling DNA in Nanopores*, Pan Stanford, 2013.
155. Anthony J. C. Ladd, *Dynamical simulations of sedimenting spheres*, Phys. Fluids A, **5**, no. 2, 299–310, 1993.
156. Stefan Kesselheim, Florian Weik, and Christian Holm, *An accurate coarse-grained DNA model for measuring current signals in DNA translocation experiments*, submitted, 2015.

Lattice-Boltzmann Simulations of Colloidal Particles at Fluid Interfaces

Jens Harting

Department of Applied Physics, Eindhoven University of Technology
P.O. Box 513, NL-5600MB Eindhoven, The Netherlands
and
Faculty of Science and Technology, Mesa+ Institute, University of Twente
P.O. Box 217, NL-7500AE Enschede, The Netherlands

E-mail: j.harting@tue.nl

Particle-stabilized fluid interfaces are very common in industrial applications as they can be found for example in the food, cosmetics, or oil industries. However, until recently our understanding of these systems was mostly based on experiments, highly simplified theoretical calculations and only very few numerical approaches. The lack of well established and widely applied simulation codes is not surprising since computer simulations of particle stabilized fluid interfaces are a highy complex task. Suitable algorithms need to be able to treat the hydrodynamics of the involved solvents, the dynamics of the suspended colloidal particles, as well as the interactions between all those constituents at the same time. Furthermore, the simulation of large scale 3D emulsions requires highly efficient and massively parallel implementations and access to state of the art supercomputers. In these lecture notes we summarize the relevant details of our own simulation method for particle stabilized interfaces which is based on a combined Lattice-Boltzmann and molecular dynamics solver. We provide an overview on important implementation details and review a number of recent applications from our group with the aim to demonstrate the specific features of particle stabilized fluid interfaces.

1 Introduction

Particle stabilized emulsions have a high potential for various purposes with industrial applications, such as cosmetics, improved low-fat food products, ice cream, drug delivery, or tertiary oil recovery[1–3]. While amphiphilic surfactant molecules are traditionally employed as emulsification agents, their effects can be mimicked or supplemented by the use of colloidal particles. These may be a cheaper or less toxic alternative to surfactants, but most importantly they may be customized to include additional desirable properties. Examples include ferromagnetic particles[4,5], particles with different interfacial properties on different parts of their surface (for example Janus particles)[6–10], or nonspherical particles[11–18], where the geometric anisotropy has an impact on their stabilization properties. Those chemically or geometrically anisotropic particles will under certain conditions deform the surrounding fluid-fluid interface. This leads to capillary interactions which can be used to tune the attraction or repulsion of adsorbed colloidal particles and as such might find applications such as the formation of new soft and highly tunable materials[17,10].

The energy differences involved in the adsorption of colloidal particles at a fluid interface are generally orders of magnitude larger than thermal fluctuations. Therefore, this adsorption process is practically irreversible[19,18] and Ostwald ripening can be fully blocked[20–24]. In this manner, particles allow for long-term stabilization of an emulsion[15,25].

However, colloidal particles stabilize fluid interfaces kinetically and not thermodynamically. They reduce the interfacial free energy by their presence: maintaining a fluid-fluid

interface is energetically expensive, and it is favourable to replace it with particle-fluid interface. This is exactly what happens when a particle adsorps to such a fluid-fluid interface. In contrast, due to their amphiphilic ineractions with the involved fluids, surfactants reduce the interfacial tension directly, which also reduces the total interfacial free energy. These effects are described in some detail in Frijters at al.[26]. Note that emulsions are in general not necessarily thermodynamically stable, as the energy cost of retaining the remaining interfacial area is still larger than the gain in entropy when the fluids mix completely. There are several ways to improve the stabilization properties, for example by using anisotropic particles with an additional rotational degree of freedom[15].

Particle-stabilized emulsions can present in various forms, classified by the shapes and sizes of their fluid domains. The form that most resembles a traditional, surfactant-stabilized emulsion is the "Pickering emulsion"[27,28], which consists of particle-covered droplets of one fluid suspended in another fluid. A more recent discovery is the bicontinuous interfacially jammed emulsion gel, or "bijel", which is characterized by two large continuous fluid domains that are intertwined and are only stable because of the particles present at their interfaces. This form was first predicted by numerical simulations[29], and shortly thereafter confirmed experimentally[30,31]. Parameters that affect the final state of an emulsion include the ratio between the two fluid components, the volume fraction of the particles, and their wettability. These parameters have been studied numerically by various authors[15,25,32–35].

Computer simulations of particle stabilized emulsions require a solver for the dynamics of the involved fluid species combined with an algorithm to track the suspended particles and their interactions. Here we use the Lattice-Boltzmann (LB) method for the fluid components, coupled to solid particles whose inter-particle interactions are simulated by molecular dynamics. This approach has proven very successful during recent years due to its ease of implementing multiphase flows, inherent parallelism allowing to harness the power of state of the art supercomputers, and straightforward methods to couple the fluid- and particle solvers. The simulation method is briefly explained in the following section 2. Within the scope of this lecture we then review three distinct applications of the method where we studied fundamental properties of particle stabilized fluid interfaces before we conclude in Sec. 4. The current lecture nodes represent a shortened collection of several original articles[36,26,37,38,18,15,25] and the interested reader is referred to those for more details.

2 Simulation Method

2.1 The Lattice-Boltzmann method

For the simulation of the fluids we apply the Lattice-Boltzmann method which is based on the discrete form of the Boltzmann equation:[39]

$$f_i^c(\mathbf{x} + \mathbf{c}_i\Delta t, t + \Delta t) = f_i^c(\mathbf{x}, t) + \Omega_i^c(\mathbf{x}, t), \tag{1}$$

where $f_i^c(\mathbf{x}, t)$ is the single-particle distribution function for a fluid component c with discrete lattice velocity $\mathbf{c}_i$ at time t located at lattice position $\mathbf{x}$. Numerous discretizations in space and time are possible, but the most popular one in three dimensions is the so-called

D3Q19 lattice. The distance between lattice nodes is given by the lattice constant Δx and nineteen directions are allowed for the velocity. Δt is the timestep and

$$\Omega_i^c(\mathbf{x},t) = -\frac{f_i^c(\mathbf{x},t) - f_i^{\mathrm{eq}}(\rho^c(\mathbf{x},t),\mathbf{u}^c(\mathbf{x},t))}{(\tau^c/\Delta t)} \tag{2}$$

is the Bhatnagar-Gross-Krook (BGK) collision operator[40]. The fluid density is defined as

$$\rho^c(\mathbf{x},t) = \rho_0 \sum_i f_i^c(\mathbf{x},t), \tag{3}$$

where ρ_0 is a unit mass factor. τ^c denotes the relaxation time for component c and

$$f_i^{\mathrm{eq}}(\rho^c,\mathbf{u}^c) = \zeta_i \rho^c \left[1 + \frac{\mathbf{c}_i\cdot\mathbf{u}^c}{c_s^2} + \frac{(\mathbf{c}_i\cdot\mathbf{u}^c)^2}{2c_s^4} - \frac{(\mathbf{u}^c\cdot\mathbf{u}^c)}{2c_s^2} + \frac{(\mathbf{c}_i\cdot\mathbf{u}^c)^3}{6c_s^6} - \frac{(\mathbf{u}^c\cdot\mathbf{u}^c)\,(\mathbf{c}_i\cdot\mathbf{u}^c)}{2c_s^4}\right] \tag{4}$$

is a discretized third order expansion of the Maxwell-Boltzmann distribution function.

$$c_s = \frac{1}{\sqrt{3}}\frac{\Delta x}{\Delta t} \tag{5}$$

is the speed of sound,

$$\mathbf{u}^c = \sum_i f_i^c(\mathbf{x},t)\mathbf{c}_i / \rho^c(\mathbf{x},t) \tag{6}$$

is the fluid velocity and ζ_i denotes a coefficient depending on the direction: $\zeta_0 = 1/3$ for the zero velocity, $\zeta_{1,\ldots,6} = 1/18$ for the six nearest neighbors and $\zeta_{7,\ldots,18} = 1/36$ for the next nearest neighbors in diagonal direction. The kinematic viscosity can be calculated as

$$\nu^c = c_s^2 \Delta t \left(\frac{\tau^c}{\Delta t} - \frac{1}{2}\right). \tag{7}$$

In the following we choose $\Delta x = \Delta t = \rho_0 = 1$ for simplicity, but a conversion to SI units is trivial[41]. In all simulations the relaxation time is set to $\tau^c \equiv 1$.

2.2 A multicomponent Lattice-Boltzmann method

Several extensions for the Lattice-Boltzmann method have been developed to simulate multicomponent and multiphase fluids[42–46]. The article Liu et al.[47] provides an extensive overview. Here, we use the multicomponent pseudopotential method introduced by Shan and Chen[42]. Every species has its own set of distribution functions following Eq. (1). A mean field force

$$\mathbf{F}^c(\mathbf{x},t) = -\Psi^c(\mathbf{x},t) \sum_{c'} g_{cc'} \sum_{\mathbf{x}'} \Psi^{c'}(\mathbf{x}',t)(\mathbf{x}'-\mathbf{x}) \tag{8}$$

is calculated locally and acts between the components. The summation includes the different fluid species c' and $\mathbf{x}'$, the nearest neighbors of lattice positions $\mathbf{x}$. $g_{cc'}$ is a phenomenological coupling constant between the species and $\Psi^c(\mathbf{x},t)$ is a monotonous weight function representing an effective mass. For the results presented here, the form

$$\Psi^c(\mathbf{x},t) \equiv \Psi(\rho^c(\mathbf{x},t)) = 1 - e^{-\rho^c(\mathbf{x},t)} \tag{9}$$

is used. To incorporate $\mathbf{F}^c(\mathbf{x},t)$ in f_i^{eq} we define

$$\Delta\mathbf{u}^c(\mathbf{x},t) = \frac{\tau^c\mathbf{F}^c(\mathbf{x},t)}{\rho^c(\mathbf{x},t)}. \tag{10}$$

The macroscopic velocity included in f_i^{eq} is shifted by $\Delta\mathbf{u}^c$ as

$$\mathbf{u}^c(\mathbf{x},t) = \frac{\sum_i f_i^c(\mathbf{x},t)\mathbf{c}_i}{\rho^c(\mathbf{x},t)} - \Delta\mathbf{u}^c(\mathbf{x},t). \tag{11}$$

As we are interested in immiscible fluids we choose a positive value for $g_{cc'}$ which leads to a repulsive interaction and thus to the emergence of surface tension. A typical range of the coupling parameter is $0.08 \leq g_{cc'} \leq 0.14$.

2.3 Colloidal particles

We couple a traditional molecular dynamics algorithm to the Lattice-Boltzmann solver in order to compute the interactions and trajectories of the suspended particles. The particle trajectories follow Newton's equations of motion

$$\mathbf{F} = m\dot{\mathbf{u}}_{\mathrm{par}}, \quad \mathbf{D} = J\dot{\omega}_{\mathrm{par}}, \tag{12}$$

which are integrated using a classical leap frog integrator. $\mathbf{F}$ and $\mathbf{D}$ are the force and torque acting on the particle with mass m and moment of inertia J. $\mathbf{u}_{\mathrm{par}}$ and ω_{par} are the velocity and the rotation vector of the particle.

The particles are discretized on the Lattice-Boltzmann lattice. They are coupled to both fluid species by a modified bounce-back boundary condition as pioneered by Ladd and Aidun[33,48–52]. The lattice Boltzmann equation then becomes

$$f_i^c(\mathbf{x}+\mathbf{c}_i, t+1) = f_i^c(\mathbf{x}+\mathbf{c}_i, t) + \Omega_i^c(\mathbf{x}+\mathbf{c}_i, t) + \mathcal{C}. \tag{13}$$

$\mathcal{C}$ depends linearly on the local particle velocity, $\bar{i}$ is defined in a way that $\mathbf{c}_i = -\mathbf{c}_{\bar{i}}$ is fulfilled. A change of the fluid momentum due to a particle leads to a change of the particle momentum in order to keep the total momentum conserved:

$$\mathbf{F}(\mathbf{x},t) = \big(2f_{\bar{i}}^c(\mathbf{x}+\mathbf{c}_i,t) + \mathcal{C}\big)\mathbf{c}_{\bar{i}}. \tag{14}$$

If the particle moves, some lattice nodes become free and others become occupied. The fluid on the newly occupied nodes is deleted and its momentum is transferred to the particle as

$$\mathbf{F}(\mathbf{x},t) = -\sum_c \rho^c(\mathbf{x},t)\mathbf{u}^c(\mathbf{x},t). \tag{15}$$

A newly freed node (located at $\mathbf{x}$) is filled with the average density of the N_{FN} neighboring fluid lattice nodes $\mathbf{x}_{i_{\mathrm{FN}}}$ for each component c[33],

$$\overline{\rho}^c(\mathbf{x},t) \equiv \frac{1}{N_{\mathrm{FN}}}\sum_{i_{\mathrm{FN}}} \rho^c(\mathbf{x}+\mathbf{c}_{i_{\mathrm{FN}}},t). \tag{16}$$

Hydrodynamics leads to a lubrication force between the particles. This force is reproduced automatically by the simulation for sufficiently large particle separations. If the distance between the particles so small that there is no lattice node between them anymore, this is

supposed to fail. Therefore, if the smallest distance between two identical spheres with radius R is smaller than a critical value $\Delta_c = \frac{2}{3}$ a lubrication correction is introduced as[52]

$$\mathbf{F}_{ij} = \frac{3\pi\mu R^2}{2}\hat{\mathbf{r}}_{ij}(\hat{\mathbf{r}}_{ij}(\mathbf{u}_i - \mathbf{u}_j))\left(\frac{1}{r_{ij} - 2R} - \frac{1}{\Delta_c}\right). \tag{17}$$

Here, μ is the dynamic viscosity, $\hat{\mathbf{r}}_{ij}$ a unit vector pointing from one particle center to the other one and $\mathbf{u}_i$ is the velocity of particle i. These corrections can be generalized to nonspherical particles in several ways. A simple, but less accurate approach is based on a scaling of the potentials following Berne and Pechukas[53,54,35]. By appropriately determining the distance between particle surfaces and also including tangential forces, the accuracy can be improved substantially[55]. For very small particle distances, the diverging nature of the lubrication correction can cause the simulation to become unstable. Since this case only happens a very few times even in very dense systems, we introduce a Hertz potential[56] which has the following shape for two identical spheres with radius R:

$$\phi_H = K_H(2r_p - r)^{5/2} \text{ for } r < 2r_p. \tag{18}$$

Here, r is the distance between particle centers. For larger distances ϕ_H vanishes. K_H is a force constant and is chosen to be $K_H = 100$ in all our simulations.

The Shan-Chen forces also act between a node in the outer shell of a particle and its neighboring node outside of the particle. This leads to an unphysical increase of the fluid density around the particle. In order to avoid this effect, the lattice nodes in the outer shell of the particle are filled with a virtual fluid with a fluid density corresponding to the average of the value in the neighboring free nodes for each fluid component:

$$\rho^c_{\text{virt}}(\mathbf{x}, t) = \overline{\rho}^c(\mathbf{x}, t). \tag{19}$$

This can be used to control the wettability of the particle surface. We define the parameter $\Delta\rho$ named "particle color". Positive values of $\Delta\rho$ are added to the "red" fluid component,

$$\rho^r_{\text{virt}} = \overline{\rho}^r + \Delta\rho, \tag{20}$$

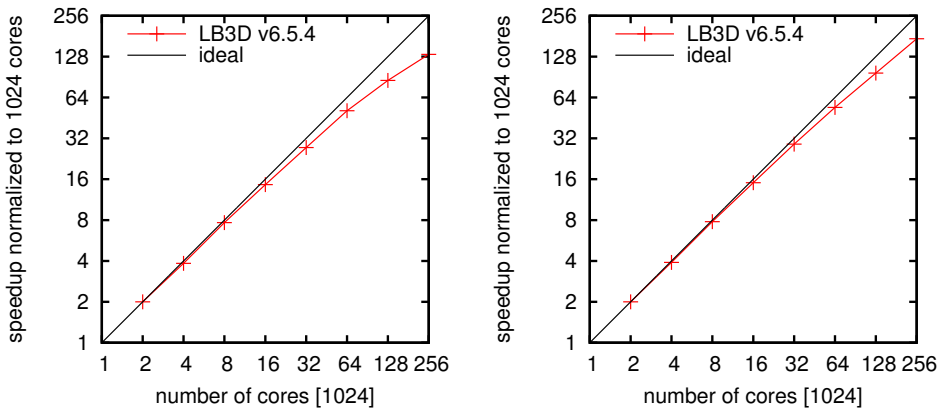
and negative values are added to the "blue" component,

$$\rho^b_{\text{virt}} = \overline{\rho}^b + |\Delta\rho|. \tag{21}$$

This leads to an approximately linear relation between $\Delta\rho$ and the three-phase contact angle θ_p[35].

2.4 Implementation

The locality of the Lattice-Boltzmann algorithm makes the implementation of massively parallel codes straightforward[57]. Generally, a regular, orthogonal grid is used, and the collision operator and boundary implementations are based on local operations so that at each lattice node only information from its own location is required. The computationally most demanding parts of a typical Lattice-Boltzmann code are "streaming" and "collision" routines. In practice, the number of floating point operations required at every timestep is not the limiting factor, but the random access in memory required for the streaming and eventual non-local force computations makes the algorithm a memory-bound numerical method. This implies a number of issues for the implementation of highly efficient codes,



(a) (b)

Figure 1: Strong scaling of LB3D on JUQUEEN. (a) relates to a system with only one fluid component, (b) to one with two fluid species and suspended particles[37,35].

be it on shared-memory multicore nodes, distributed memory clusters, or accelerator cards such as graphical processing units (GPUs) or Xeon Phi cards.

Several well established and highly scalable Lattice-Boltzmann implementations exist, which have all demonstrated excellent scaling on hundreds of thousand CPUs. To name a few, *Ludwig*[58], *LB3D*[59], *walBerla*[60], *MUPHY*[61], and *Taxila* LBM[62] are all codes which have been used for numerous high quality scientific publications. Most of these codes are able to handle solid objects suspended in fluids. The first three can even combine this with multiple fluid components or phases.

Our own implementation, *LB3D*, is written in Fortran 90 and parallelized using MPI. Long-running simulations on massively parallel architectures require parallel I/O strategies and checkpoint and restart facilities. *LB3D* uses the parallel HDF5 library for I/O which has proven to be highly robust and performant on all supercomputers we had access to. *LB3D* has been shown to scale almost linearly on up to 262,144 cores on the European Blue Gene systems Jugene and Juqueen based at the Jülich Supercomputing Centre in Germany[36,47]. However, such excellent scaling required some optimizations of the code: Initially, LB3D showed only low efficiency in strong scaling beyond $65\,536$, which could be related to a mismatch of the network topology of the domain decomposition in the code and the network actually employed for point-to-point communication. The Blue Gene/P provides direct links only between direct neighbors in a three-dimensional torus, so a mismatch can cause severe performance losses. Allowing MPI to reorder process ranks and manually choose a domain decomposition based on the known hardware topology, efficiency can be brought close to ideal. The importance of these implementation details is depicted by strong scaling measurements based on a system of $1\,024^2 \times 2\,048$ lattice sites carrying only one fluid species (Fig. 1a) and a similarly sized system containing two fluid species and $4\,112\,895$ uniformly distributed particles with a diameter of ten lattice units (Fig. 1b).

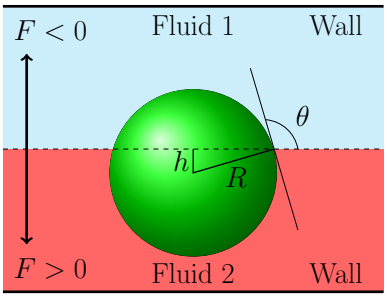


Figure 2: Equilibrium state of a spherical particle at a fluid-fluid interface. The contact angle, $\theta = \cos^{-1}(h/R)$, where h is the height of the particle centre of mass above the interface and R the particle radius, is defined with respect to Fluid 1. The force, F, on the particle acts to detach it into the wetting (positive force) or non-wetting (negative force) fluid (Figure reprinted from Davies et al.[18]).

3 Applications

3.1 Detachment energies of spheroidal particles from fluid-fluid interfaces

The detachment energy of a single particle from a fluid-fluid interface plays a crucial role in our understanding of particle-stabilised emulsions. Several authors have studied the detachment of particles in the past, but to the best of our knowledge, none of these extended their treatment to the case of anisotropic spheroidal particles, which we focused on in a recent paper by Davies et al.[18]. This section gives a summary of the results presented in that publication.

Previous detachment energy studies focussing on free energy differences between an equilibrated particle at an interface[63–70] and in the bulk revealed a crucial dependence on particle shape: prolate and oblate spheroidal particles attach to interfaces more strongly because they reduce the interface area more than spherical particles for a given particle volume.[71–75] For a particle already adsorbed at an interface to detach itself, the particle must deform the interface and overcome the interface's resistive force: there is a free-energy barrier and an associated activation energy.

We develop a simple thermodynamic model for the detachment energy of spheroidal particles from fluid-fluid interfaces as a function of contact angle and aspect ratio only, and highlight the implications of our simplifications. Assuming that the effect of gravity can be neglected, the surface free energy of a particle at an interface (Fig. 2) is given by

$$E = \sigma_{12}A_{12} + \sigma_{p1}A_{p1} + \sigma_{p2}A_{p2}, \tag{22}$$

where A_{ij} is the area and σ_{ij} is the surface-energy of the i,j interface where $i,j = \{1\text{: fluid 1, 2: fluid 2, p: particle}\}$.[76] We neglect line-tension since it is relevant only for nano-sized particles.[72] The surface area of the particle is $A_p = A_{p1} + A_{p2}$. The free energy of a system in which the particle is fully immersed in either fluid 1 or fluid 2 is given by $E_i = \sigma_{12}A_{12} + \sigma_{pi}A_{pi}$ where $i = 1, 2$. Taking the free energy difference between a spherical particle at an interface (Fig. 2) and a spherical particle immersed in the bulk fluid yields the detachment energy[73,74]

$$E = \pi R^2 \sigma_{12}(1 - |\cos\theta|)^2. \tag{23}$$

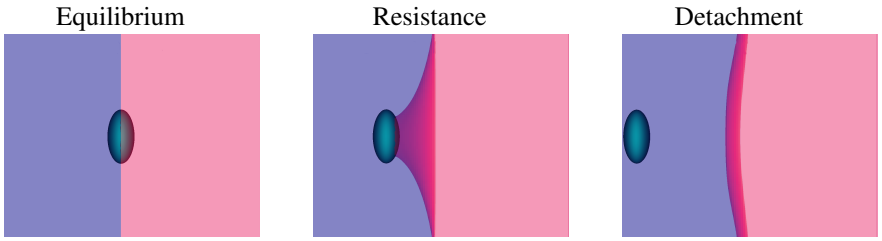


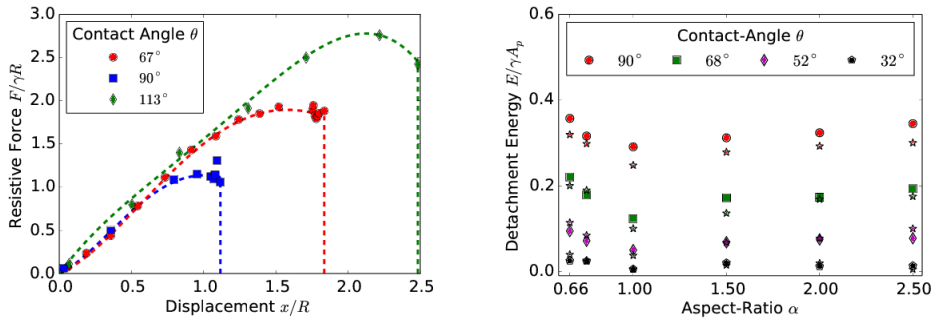
Figure 3: Snapshots of a prolate spheroidal particle of aspect ratio $\alpha = 2$ under the influence of an external force detaching from an interface. For each snapshot, we run a simulation with the particle fixed and measure the resistive force from the interface on the particle (Figures reprinted from Davies et al.[18]).

For neutrally wetting micron-sized particles at an interface with surface-tension $\sigma_{12} = 50\,\mathrm{mN\,m^{-1}}$, the detachment energy is much larger than the thermal energy, $E/k_BT \sim 10^7$, and particles irreversibly attach to the interface. For nano-sized particles at the same interface with very large or small contact angles, $E \sim k_BT$, and particles may freely adsorb to and desorb from the interface[19].

For our simulations, we initialise a system volume of size 128^3 in lattice units which is half filled with liquid 1 and half filled with liquid 2 ($\rho^{(1)} = \rho^{(2)} = 0.7$), such that an interface forms at $x = 64$. The top and bottom are closed with solid walls. The particle density is chosen as $\rho_p = 2$ which is an arbitrary choice. The particle is placed at the interface and is not under the influence of any external forces. We first equilibrate the system until the interface diffuses and the particle establishes its equilibrium position, and hence its contact angle, on the interface.[33] Then, we apply a constant external force to the particle. As stated above, the wettability of the particles can be tuned and we determine the contact angle by subtracting the height of the particle centre of mass above the interface (we linearly interpolate the interface position) and dividing by the particle radius, $\cos\theta = \frac{h}{R}$.[33,26]

To obtain the minimum detachment force, we employ a binary search algorithm: we start the algorithm by using the fact that the particle remains attached at the interface for a zero-force, $F_{\mathrm{att}} = 0$, and guessing a force which detaches the particle, F_{det}. We then run a new simulation with a force, $F_{\mathrm{new}} = \frac{1}{2}(F_{\mathrm{att}} + F_{\mathrm{det}})$ and repeat this procedure until we determine F_{det} to the desired accuracy. As a next step, we run a single simulation with the minimum detachment force, saving the simulation state frequently. We then run several simulations from the saved simulation snapshots (Fig. 3) but now with the particles fixed so that drag, buoyancy and gravity forces can be neglected. We let the systems from each snapshot equilibrate, and we measure the resulting force on the particle which is exactly the resistive force of the interface. This allows us to build a force-distance curve $F(x)$. We fit $F(x)$ with a fourth-order polynomial, which allows us to capture the linear regime and the detachment break-off regime accurately, and integrate the fitted function numerically to obtain the detachment energy. The detachment distance is the minimum distance at which the resistive force exerted by the interface on the particle is zero. As discussed shortly, the resistive force decreases discontinuously to zero at the point of detachment.

The restoring force provided by the interface to the particle as a function of displacement from equilibrium is shown in Fig. 4a. The corresponding fourth-order polynomial



(a) (b)

Figure 4: a) The normalized resistive force for a spherical particle of radius $R = 10$ is approximately linear for small displacements as predicted by de Gennes et al.[77,78] and O'Brien.[79] The symbols are simulation data and the dashed lines represent fourth-order polynomial fits to that data. The fourth-order fits are integrated in order to obtain the detachment energy, E
b) Dependence of the detachment energy on the aspect ratio, α, for several different contact angles, θ. Each set of coloured data points represents different wettabilities of the particles. Stars are theoretical calculations from our thermodynamic model in eq. (24) and eq. (25) and symbols are numerical data (Figures reprinted from Davies et al.[18]).

fits are also shown.

Davies et al.[18] introduce a simplified thermodynamic model describing the detachment energies of prolate and oblate spheroids in their equilibrium positions from interfaces.

$$\Delta E^{\perp} = \frac{\bar{h}^2}{2}\left(1 - \frac{\alpha}{2G(\alpha)}\right) - \frac{\bar{h}}{2} + \frac{\alpha}{4G(\alpha)}, \tag{24}$$

$$\Delta E^{\parallel} = \frac{\bar{h}^2}{2}\left(1 - \frac{\alpha}{2G(\alpha)}\right) - \frac{\bar{h}}{2} + \frac{1}{4G(\alpha)}. \tag{25}$$

These approximate expressions are simple quadratic functions of the dimensionless height $\bar{h}$ and the aspect ratio α. $G(\alpha)$ is a geometry factor. In Fig. 4b we compare the analytical results from Eq. (24) and Eq. (25) with our simulation data.

We find good agreement between our thermodynamic model and the numerical simulations. For neutrally wetting particles, the measured detachment energies from simulations are larger than those predicted by the thermodynamic model; this is expected since the particle has to deform the interface to overcome its resistive force. The differences are of the order of 10%, suggesting that the thermodynamic model, which does not take into account interface deformations, is fairly accurate for the particle aspect ratios we investigated. Similarly, we find good agreement for $\theta = 68°$ where the numerical data show a higher detachment energy than predicted by the thermodynamic model, as expected. For contact angles $\theta = 52°$ and $\theta = 32°$, we still find good qualitative agreement between thermodynamic theory and numerical simulations for both prolate and oblate spheroids. However, for oblate spheroids the numerical detachment energy is less than the analytical predictions, though within errors of the order of the symbol size.

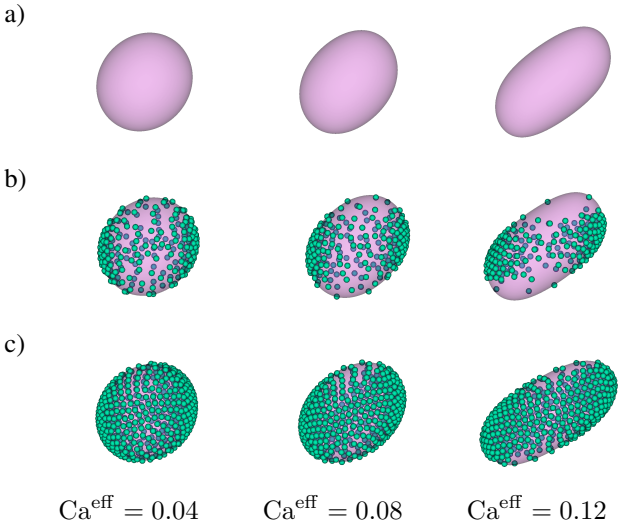


Figure 5: Side-view of deformed droplets (a) $\chi = 0.00$, b) $\chi = 0.27$ and c) $\chi = 0.55$). The particles prefer to stay in the center of the channel where the shear flow is weakest. The particles further exhibit tank-treading-like behaviour: they move around the interface following the shear flow (Figure reprinted from Frijters et al.[26]).

3.2 Shear induced deformation of particle-armored droplets

Droplets covered by colloidal particles appear as a building block of emulsions, or might be produced as individual containers for example for drug delivery purposes. In this section we review some simulation results on how such droplets behave under an externally imposed shear flow. The section is based on an article by Frijters et al.[26] and the reader is referred to this article for more details.

A fraction χ of the interface of a droplet immersed in a second fluid is covered with particles and a shear flow is imposed by moving boundaries at the top and bottom of the simulation domain, which causes the droplet to deform. The resulting deformation can be quantified using Taylor's dimensionless deformation parameter[80,81]

$$D \equiv \frac{L - B}{L + B}. \tag{26}$$

Here L and B are the length and the breadth of the droplet. For small deformations, Taylor predicts a linear dependence of the deformation of a droplet on the capillary number (see Eq. 27), which becomes $D = 35/32\mathrm{Ca}$ for equiviscous fluids. Due to the finite system size, we define the capillary and Reynolds numbers in terms of an effective shear rate $\dot{\gamma}^{\mathrm{eff}}$, which is measured at the interface of the droplet during the simulation:

$$\mathrm{Ca}^{\mathrm{eff}} \equiv \frac{\mu_m \dot{\gamma}^{\mathrm{eff}} R_d}{\sigma}, \qquad \mathrm{Re}^{\mathrm{eff}} \equiv \frac{\rho_m \dot{\gamma}^{\mathrm{eff}} R_d^2}{\mu_m}, \tag{27}$$

These are defined in terms os the dynamic viscosity μ_m of the medium, the radius of the initial undeformed droplet R_d and the surface tension σ. The effective Reynolds number is varied between approximately $0.6 < \mathrm{Re}^{\mathrm{eff}} < 25$.

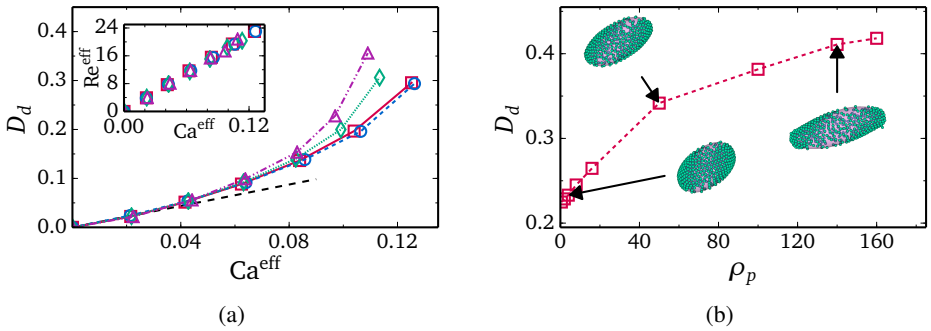


Figure 6: a) The deformation parameter D as a function of the effective capillary number $\mathrm{Ca}^{\mathrm{eff}}$ for $\chi = 0$ (red squares), $\chi = 0.27$ (blue circles), $\chi = 0.41$ (green diamonds), and $\chi = 0.55$ (purple triangles). The effect of adsorped particles is very weak for low χ, becomes noticeable at $\chi > 0.4$. Taylor's law is reproduced for small Ca (dashed line). The inset depicts that the Reynolds number scales linearly with the capillary number. b) The deformation parameter D as a function of the rescaled mass of the particles m_p^*, with $\chi = 0.55$ and $\mathrm{Ca}^{\mathrm{eff}} = 0.1$. The inertia of the heavier particles causes additional deformation as they drag the droplet interface in the direction of the shear flow (Figures reprinted from Frijters et al.[26]).

Let us first understand the effect of nanoparticles on the deformation properties of the droplet, by discussing how they position themselves at and move over the droplet interface. The fluid-fluid interaction strength is held fixed at $g_{br} = 0.10$ and the particles with radius $r_p = 5.0$ are neutrally wetting ($\theta_p = 90°$) and have a mass density which is identical to the fluid mass density. The simulation volume is chosen to be $n_x = n_y = 256$, $n_z = 512$, with an initial droplet radius of $R_d^{\mathrm{init}} = 0.3 \cdot n_x = 76.8$, while the number of particles is varied as $n_p = 0$, 128, 256, 320, 384, 446 and 512. This results in surface coverage fractions between $\chi = 0$ and $\chi = 0.55$. The capillary number is changed by changing the shear rate. Some examples of the deformations thus realised are shown in Fig. 5, for $\mathrm{Ca}^{\mathrm{eff}} = 0.04$, 0.08, 0.12 and $\chi = 0.0$ (a), $\chi = 0.27$ (b) and $\chi = 0.55$ (c).

The interface of a sheared and thus deformed droplet is increased as compared to its original spherical shape and more space is available for the particles to move freely over the interface (cf. Fig. 5). The particles are swept over the interface with increasing velocity as they move away from the center plane of the system and up the shear gradient. If the particles would not be affected by the shear flow, they would prefer to occupy interface with high local curvature as can be explained by a geometrical argument: the interface removed by a spherical particle at a curved interface is larger than the circular area removed from a flat interface. This explains why in this dynamic equilibrium, most particles can be found at the tips of the droplet (see Fig. 5 b)). Even though the overall structure of the particles on the droplet interface remains stable over time, individual particles move over the interface, performing a quasi-periodic motion.

The linear dependence of the deformation parameter D on the effective capillary number introduced above is recovered for low capillary number and low particle coverage (cf. Fig. 6a). When the coverage fraction grows beyond $\chi > 0.40$ the deformations in this regime increase with increasing χ and constant capillary number. Combining Eqns. 27

results in a relation between the capillary and Reynolds number:

$$\mathrm{Re}^{\mathrm{eff}} = \sigma \left(\frac{\rho_m R_d}{\mu_m^2} \right) \mathrm{Ca}^{\mathrm{eff}} \tag{28}$$

Since we change $\mathrm{Ca}^{\mathrm{eff}}$ explicitly by changing the shear rate, $\mathrm{Re}^{\mathrm{eff}}$ is proportional to the capillary number for a fixed value of the surface tension and inertial effects increase the deformation. Furthermore, since the particles do not affect the surface tension, all curves of the Reynolds number versus capillary number have the same slope (cf. inset of Fig. 6a and Eq. (28)). This implies that the increased deformation in the case of added particles is not caused by changes in inertia of the fluids, but instead the inertia of the particles themselves plays a decisive role here: As shown in Fig. 6b, we have varied the mass of the particles over two orders of magnitude. We kept $\chi = 0.55$ and $\mathrm{Ca}^{\mathrm{eff}} = 0.1$ constant and rescaled the mass scale with the reference mass: $m_p^* = m_p/524$. The particles are accelerated as long as they are on the part of the droplet interface that experiences a shear flow at least partially parallel to the particle movement. Eventually, particles have to "round the corner". The increased inertia of heavier particles makes it more difficult to change their movement, leading to a situation where the droplet interface is in fact initially dragged farther away in the direction of the shear flow instead. This then explains the increase of deformation with increasing particle mass. As our deformation is increased substantially, the system size limits the deformation we can induce. Therefore, the values presented here are underpredictions of the actual effect of increased mass at high deformations.

3.3 Timescales of emulsion formation caused by anisotropic particles

In this section different types of particle stabilized emulsions and the effect of the particle shape on some of their properties are discussed. The results presented here are a summary of an article recently published by Günther et al.[35].

We find two different types of emulsions in our simulations, namely the Pickering emulsion (Fig. 7, left) and the bijel (Fig. 7, right). The choice of parameters (such as particle contact angle, particle concentration, fluid-fluid ratio, particle aspect ratio) determines the type of emulsions. Parameter studies for emulsions have been discussed in the past.[33,35,25] for spherical and ellipsoidal particles, respectively. Here, we limit ourselves to anisotropy effects on the time dependence of the emulsion formation. We use prolate ellipsoids ($m = 2$; Fig. 7, top), spheres ($m = 1$; Fig. 7, center) and oblate ellipsoids ($m = 1/2$; Fig. 7, bottom). The interaction parameter between the two fluids (see Eq. (8)) is chosen as $g_{br} = 0.08$ which corresponds to a fluid-fluid interfacial tension of 0.0138. The particles are neutrally wetting (contact angle $\theta_p = 90°$) and the particle volume concentration is chosen as $\Xi = 0.24$. The simulated systems have periodic boundary conditions in all three directions and a side length of $L_S = 256\Delta x$. Initially, the particles are distributed randomly. At each lattice node a random value for each fluid component is chosen so that the designed fluid-fluid ratio is kept (1:1 for the bijels and 5:2 for the Pickering emulsions). When the simulation evolves in time, the fluids separate and droplets/domains with a majority of one of the fluids form.

The average size of droplets/domains $L(t)$ can be determined by measuring

$$L(t) = \frac{1}{3} \sum_{i=x,y,z} L(t)_i, \text{ with } \quad L(t)_i = \frac{2\pi}{\sqrt{\langle k_i^2(t) \rangle}}. \tag{29}$$

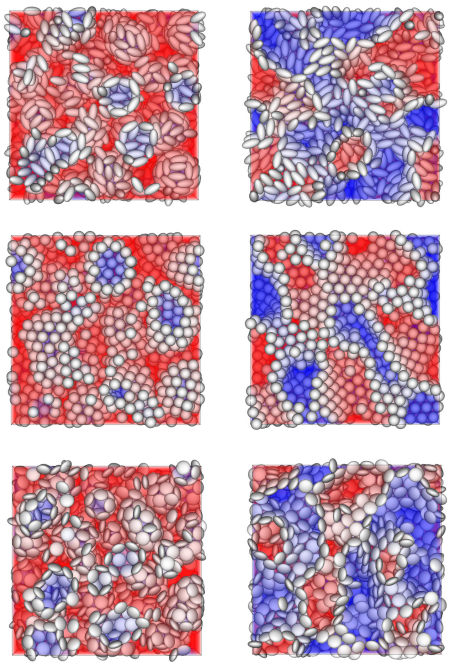
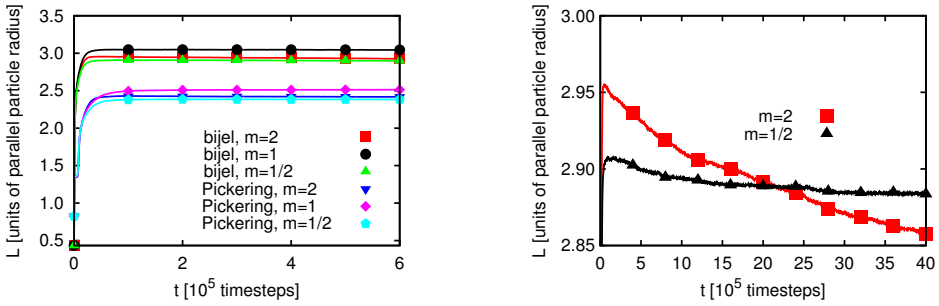


Figure 7: Snapshots of typical simulated Pickering emulsions (left) and bijels (right) after 10^5 timesteps. The emulsions are stabilized by prolate ellipsoids ($m = 2$, top), spheres ($m = 1$, center) and oblate ellipsoids ($m = 1/2$, bottom). The parameter determining if one obtains a bijel or a Pickering emulsion is the fluid ratio which is chosen as 1:1 for the bijels and 5:2 for Pickering emulsions (Figure reprinted from Günther et al.[15]).

$\langle k_i^2(t)\rangle = \sum_{\mathbf{k}} k_i^2(t)\varsigma(\mathbf{k},t)/\sum_{\mathbf{k}} k_i^2(t)$ is the second-order moment of the three-dimensional structure function $\varsigma(\mathbf{k},t) = (1/\varsigma_n)|\varphi'_{\mathbf{k}}(t)|$. $\varphi' = \tilde{\varphi} - \langle\tilde{\varphi}\rangle$ is the fluctuation of $\tilde{\varphi}$ which is the Fourier transform of the order parameter field $\varphi = \rho^r - \rho^b$.

The time development of $L(t)$ for the three different particle types (prolate, spherical and oblate ($m = 2$, 1 and $1/2$) and for Pickering emulsions and bijels is shown in Fig. 8a. We can identify three regimes: in the first few hundred timesteps the initial formation of the droplets/domains starts. Then, the growth of droplets/domains is being driven by Ostwald ripening. At even later times, droplets/domains grow due to coalescence. When two droplets unify, the area coverage fraction of the particles at the interface is increased because the surface area of the new droplet is smaller than that of the two smaller droplets before. At some point the area coverage fraction of the particles is sufficiently high to prevent further coalescence. The state which is reached at that time is (at least kinetically) stabilized and one obtains a stable emulsion. The values for $L(t)$ are larger for bijels than for Pickering emulsions. This can be explained by the way we calculate $L(t)$ (see Eq. (29) and related text) using a Fourier transformation of the order parameter field.

It can clearly be seen that anisotropic particles are more efficient in interface stabilization than spheres since they can cover larger interfacial areas leading to smaller



(a) (b)

Figure 8: a) Pickering emulsion and bijel: Time development of the average domain size $L(t)$ (see Eq. (29)) for $m = 1$ and $m = 2$. At first view, a steady state is reached after about 10^5 timesteps. $L(t)$ is larger for bijels than for Pickering emulsions, which is due to the measurement being based on the Fourier transform of the order parameter. Ellipsoids are able to stabilize larger interface areas than spheres leading to smaller $L(t)$.
b) Bijel: Zoom for $m = 2$ and $m = 1/2$: Time development of the average domain sizes depicting the impact of the additional timescales. The range of the variation of $L(t)$ is larger as compared to the Pickering emulsion due to the impact of a small deformation on the larger effective interface of the bijel (Figures reprinted from Günther et al.[15]).

fluid domains (note that the simulation volume is kept constant). However, the difference in $L(t)$ for $m = 2$ and $m = 1/2$ is small. This can be understood as follows: if a neutrally wetting prolate ellipsoid is adsorbed at a flat interface, it occupies an area $A_{P,F}(m > 1) = m^{1/3} A_{p,s}$, where $A_{p,s}$ is the occupied interface area for a sphere with the same volume. This corresponds in the case of $m = 2$ to the occupied interface being larger by a factor of 1.26 as compared to spheres. For an oblate ellipsoid the occupied interface area is $A_{P,F}(m < 1) = m^{-2/3} A_{p,s}$ which for $m = 1/2$ is by a factor of 1.59 larger than the area occupied by spheres. Since in emulsions the interfaces are generally not flat, these formulae can only provide a qualitative explanation of the behavior of $L(t)$: If the interface curvature is not neglectable anymore, we lose some of the efficiency of interface stabilization, which is more pronounced for $m < 1$. This explains why the value of $L(t)$ for $m = 1/2$ is only slightly smaller than for $m = 2$.

It seems that $L(t)$ reaches a steady state after some 10^5 timesteps for both types of emulsions and for all three values of m. However, if one zooms in one can observe that $L(t)$ develops for a longer time period if the particles have a non-spherical shape. As will be demonstrated below, the reason for this phenomenon is the additional rotational degrees of freedom due to the particle anisotropy. Furthermore, the time development of $L(t)$ for emulsions stabilized by prolate particles requires more time than that for the oblate ones. If a particle changes its orientation as compared to the interface or a neighboring particle this generally changes the interface shape. In this way the domain sizes are influenced, leading to changes of $L(t)$ – an effect which is not observed for $m = 1$.

Fig. 8b depicts a zoom-in of the time development of $L(t)$ for bijels with $m = 2$ and $m = 1/2$, respectively. One observes that $L(t)$ decays in both cases. The range of the decay is larger for $m = 2$ than for $m = 1/2$. The time of reordering is much shorter for $m = 1/2$ as compared to $m = 2$. These effects can be explained by the presence of

additional rotational degrees of freedom for the anisotropic particles. While oblate particles have only a single additional rotational degree of freedom as compared to spheres, prolate particles show an even more complex behavior due to their second additional rotational degree of freedom.

Th first additional timescale we find can be related to the adsorption of single particles and their further rotation towards the interface during the initial stages of emulsion formations. However, while the decrease of $L(t)$ in Fig. 8b is certainly of the order of 10^6 LB timesteps, we find that the single particle dynamics can only be responsible for effects on scales of the order of 10^3 timesteps.

By means of time dependent orientational order parameters and pair correlation functions we investigate the impact of collective effects of many particles being adsorbed at a flat fluid-fluid interface. For prolate particles, here also the mutual orientation of the particles is important and one has an additional degree of freedom leading to particle orientational ordering. The presence of many particles at an interface leads to two additional timescales in the reordering. The first one is the rotation of the particle towards the interface. The particle rotates towards its final orientation parallel to the interface. For higher particle concentrations the time needed for a particle to come to its final orientation increases. Hydrodynamic as well as excluded volume effects become more important. Above a critical concentration not every particle reaches its "final" orientation. In summary, time scales involved in many-particle effects are found to be of the order of 10^4 to even several 10^5 timesteps.

In emulsions the interfaces are generally not flat. Pickering emulsions usually have (approximately) spherical droplets and a bijel has an even more complicated structure of curved interface. The simplest realization of a curved interface is a single droplet. By simulating the ordering of many prolate ellipsoids at a droplet surface, it was found that the ordering is somewhat faster than in the case of flat interfaces. This is not surprising because capillary interactions between the particles play a substantial role in the case of curved interface.

We can understand one of the additional timescales with the behavior of the ellipsoidal particles at a single droplet. The particles reorder and it can be shown that this leads to a small change of the shape of the droplet which is (almost) exactly spherical in the beginning[32]. A change of the interface shape caused by reordering of anisotropic particles leads to a change of $L(t)$. The reordering of particle ensembles at flat as well as spherical interfaces takes of the order of 10^5 timesteps. This reordering takes place in idealized systems with constant interfaces which do not change their shape considerably. In real emulsions, however, the interface geometry changes substantially during their formation. For example, two droplets of a Pickering emulsion can coalesce. After this unification the particle ordering starts anew. This explains the fact that the additional timescale we find in our emulsions is of the order of several 10^6 timesteps.

This reordering is pronounced in the case of curved interfaces, where the movement of the particles leads to interface deformations and capillary interactions. During the formation of an emulsion, droplets might coalesce (Pickering emulsions) or domains might merge (bijels). After such an event the particles at the interface have to rearrange in order to adhere to the new interface structure. Due to this, the local reordering is practically being "restarted" leading to an overall increase of the interfacial area on a timescale of at least several 10^6 timesteps.

Our findings provide relevant insight in the dynamics of emulsion formation which is generally difficult to investigate experimentally due to the required high temporal resolution of the measurement method and limited optical transparency of the experimental system. It is well known that in general particle-stabilized emulsions are not thermodynamically stable and therefore the involved fluids will always phase separate – even if this might take several months. Anisotropic particles, however, provide properties which might allow the generation of emulsions that are stable on substantially longer timescales. This is due to the continuous reordering of the particles at liquid interfaces which leads to an increase in interfacial area and as such counteracts the thermodynamically driven reduction of interface area.

4 Conclusion

We introduced our simulation method for fluid interfaces stabilized by colloidal particles, gave some details of the implementation of a highly efficient and massively parallel simulation code and reviewed a number of recent applications.

When an interface is stabilized by colloidal particles, it is of fundamental interest to understand how reversible the adsorption of the particles is. For typical particle diameters of the orders of micrometres, the energy required to detach a particle can be of the order of several thousand $k_B T$. For nanoparticles, however, thermal energies might be of the same order as the gain in free energy due to the particles being adsorped. In Sec. 3.1 and the recent publication by Davies et al.[18], we studied this problem in detail and in particular presented a simplified thermodynamic model to estimate detachment energies of spheroidal particles. The model shows good agreement with our simulations and allows the experimentalist to do ad-hoc estimates in order to for example better understand how stable an emulsion stabilized by colloidal particles of different shapes will be.

Particles and surfactants show some important differences when used for the stabilization of fluid interfaces. As shown in Sec. 3.2 and a recent paper by Frijters et al.[26], particles do not directly change the interfacial tension, but reduce the interfacial free energy by removing interface area. Furthermore, particles are massive objects and under extreme circumstances such as high mass density and strong shear forces, the inertia of the particles plays a role. This can, for example, lead to an enhanced deformation of particle armored droplets in shear flow.

The geometrical shape of the particles used to stabilize an emulsion plays an important role for the dynamics and long-term stability of the systems. This was demonstrated in Sec. 3.3 and the recent article by Günther at al.[15], where we studied the appearance of additional timescales in the dynamics of the formation of emulsions stabilized by oblate or prolate ellipsoids. The dynamics during the adsorption process of a single particle as well as interfacial jamming or capillary interaction between particles play a crucial role. These effects cause a continuous, long-term reordering of already adsorped particles which might be utilized to counteract the inherent thermodynamic instability of these systems by improved kinetic stabilization.

Acknowledgments

We thank all the students and collaborators who have contributed to the results reviewed in this lecture and the corresponding original journal articles: Fernando Bresme, Peter V. Coveney, Gary B. Davies, Stefan Frijters, Florian Günther, Florian Janoscheck, Fabian Jansen, Badr Kaoui, Timm Krüger, and Qingguang Xie.

References

1. E. Dickinson, *Food emulsions and foams: Stabilization by particles*, Current Opinion in Colloid & Interface Science, **15**, 40, 2010.
2. D.A. Pink, M. Shajahan, and G. Razul, *Computer simulation techniques for food science and engineering: simulating atomic scale and coarse-grained models*, Food Structure, **1**, 71, 2014.
3. D.E. Tambe and M.M. Sharma, *The effect of colloidal particles on fluid-fluid interfacial properties and emulsion stability*, Advances in Colloid and Interface Science, **52**, 1, 1994.
4. S. Melle, M. Lask, and G.G. Fuller, *Pickering emulsions with controllable stability*, Langmuir, **21**, 2158, 2005.
5. E. Kim, K. Stratford, and M.E. Cates, *Bijels containing magnetic particles: a simulation study*, Langmuir, **26**, 7928, 2010.
6. B.P. Binks and P.D.I. Fletcher, *Particles adsorped at the oil-water interface: A theoretical comparison between spheres of uniform wettability and "Janus" particles*, Langmuir, **17**, 4708, 2001.
7. N. Glaser, D.J. Adams, A. Böker, and G. Krausch, *Janus particles at liquid-liquid interfaces*, Langmuir, **22**, 5227, 2006.
8. J. Aveyard, *Can Janus particles give thermodynamically stable Pickering emulsions?*, Soft Matter, **8**, 5233, 2012.
9. A. Kumar, B.J. Park, F. Tu, and D. Lee, *Amphiphilic Janus particles at fluid interfaces*, Soft Matter, **9**, 6604, 2013.
10. Q. Xie, G. B. Davies, F. Günther, and J. Harting, *Tunable dipolar capillary deformations for magnetic Janus particles at fluid-fluid interfaces*, Submitted for publication, 2015.
11. M.J.A. Hore and M. Laradji, *Prospects of nanorods as an emulsifying agent of immiscible blends*, Journal of Computational Physics, **128**, 054901, 2008.
12. W. Zhou, J. Cao, W. Liu, and S. Stoyanov, *How rigid rods self-assemble at curved surfaces*, Angewandte Chemie - International Edition, **48**, 378, 2009.
13. M. Grzelczak, J. Vermant, E.M. Furst, and L.M. Liz-Marzán, *Directed self-assembly of nanoparticles*, ACS Nano, **4**, 3591, 2010.
14. L. Botto, E.P. Lewandowski, M. Cavallaro, and K.J. Stebe, *Capillary interactions between anisotropic particles*, Soft Matter, **8**, 9957, 2012.
15. F. Günther, S. Frijters, and J. Harting, *Timescales of emulsion formation caused by anisotropic particles*, Soft Matter, **10**, 4977, 2014.
16. G.B. Davies, T. Krüger, P.V. Coveney, J. Harting, and F. Bresme, *Interface deformations affect the orientation transition of magnetic ellipsoidal particles adsorbed at fluid-fluid interfaces*, Soft Matter, **10**, 6742, 2014.

17. G.B. Davies, T. Krüger, P.V. Coveney, J. Harting, and F. Bresme, *Assembling ellipsoidal particles at fluid interfaces using switchable dipolar capillary interactions*, Advanced Materials, **26**, 6715, 2014.
18. G.B. Davies, T. Krüger, P.V. Coveney, and J. Harting, *Detachment energies of spheroidal particles from fluid-fluid Interfaces*, Journal of Chemical Physics, **141**, 154902, 2014.
19. B.P. Binks, *Particles as surfactants – similarities and differences*, Current Opinion in Colloid & Interface Science, **7**, 21, 2002.
20. W. Ostwald, *Analytische Chemie*, Engelmann, 1901.
21. R.D. Vengrenovitch, *On the Ostwald ripening theory*, Acta Metallurgica, **30**, 1079, 1982.
22. Y. Enomoto, K. Kawasaki, and M. Tokuyama, *Computer modelling of Ostwald ripening*, Acta Metallurgica, **35**, 907, 1987.
23. P.W. Voorhees, *Ostwald ripening of two-phase mixtures*, Annual Review of Materials Science, **22**, 197, 1992.
24. P. Taylor, *Ostwald ripening in emulsions*, Adv. Colloid Interface Sci., **75**, 107, 1998.
25. S. Frijters, F. Günther, and J. Harting, *Domain and droplet sizes in emulsions stabilized by colloidal particles*, Physcal Review E, **90**, 042307, 2014.
26. S. Frijters, F. Günther, and J. Harting, *Effects of nanoparticles and surfactant on droplets in shear flow*, Soft Matter, **8**, 6542, 2012.
27. W. Ramsden, *Separation of solids in the surface-layers of solutions and "suspensions"*, Proceedings of the Royal Society of London, **72**, 156, 1903.
28. S.U. Pickering, *Emulsions*, Journal of the Chemical Society, Transactions, **91**, 2001, 1907.
29. K. Stratford, R. Adhikari, I. Pagonabarraga, J.-C. Desplat, and M.E. Cates, *Colloidal jamming at interfaces: a route to fluid-bicontinuous gels*, Science, **309**, 2198, 2005.
30. E.M. Herzig, K.A. White, A.B. Schofield, W.C.K. Poon, and P.S. Clegg, *Bicontinuous emulsions stabilized solely by colloidal particles*, Nature Materials, **6**, 966, 2007.
31. P.S. Clegg, E.M. Herzig, A.B. Schofield, S.U. Egelhaaf, T.S. Horozov, B.P. Binks, M.E. Cates, and W.C.K. Poon, *Emulsification of partially miscible liquids using colloidal particles: nonspherical and extended domain structures*, Langmuir, **23**, 5984, 2007.
32. E. Kim, K. Stratford, R. Adhikari, and M.E. Cates, *Arrest of fluid demixing by nanoparticles: a computer simulation study*, Langmuir, **24**, 6549, 2008.
33. F. Jansen and J. Harting, *From bijels to Pickering emulsions: a lattice Boltzmann study*, Physcal Review E, **83**, 046707, 2011.
34. S. Aland, J. Lowengrub, and A. Voigt, *A continuum model of colloid-stabilized interfaces*, Physics of Fluids, **23**, 062103, 2011.
35. F. Günther, F. Janoschek, S. Frijters, and J. Harting, *Lattice Boltzmann simulations of anisotropic particles at liquid interfaces*, Computers and Fluids, **80**, 184, 2013.
36. D. Groen, O. Henrich, F. Janoschek, P.V. Coveney, and J. Harting, "Lattice-Boltzmann methods in fluid dynamics: Turbulence and complex colloidal fluids", in: Jülich Blue Gene/P Extreme Scaling Workshop 2011, Wolfgang Frings Bernd Mohr, (Ed.). Jülich Supercomputing Centre, 52425 Jülich, Germany, apr 2011, FZJ-JSC-IB-2011-02; http://www2.fz-juelich.de/jsc/docs/autoren2011/mohr1/.

37. J. Harting, S. Frijters, F. Janoschek, and F. Günther, *Coupled lattice Boltzmann and molecular dynamics simulations on massively parallel computers*, in: NIC Symposium 2012 – Proceedings, M. Kremer K. Binder, G. Münster, (Ed.), vol. 45 of *NIC Series*, p. 243, Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag. 2012, http://hdl.handle.net/2128/4538.
38. T. Krüger, S. Frijters, F. Günther, B. Kaoui, and J. Harting, *Numerical simulations of complex fluid-fluid interface dynamics*, European Physical Journal Special Topics, **222**, 177, 2013.
39. R. Benzi, S. Succi, and M. Vergassola, *The lattice Boltzmann equation: theory and applications*, Physics Reports, **222**, 145, 1992.
40. P.L. Bhatnagar, E.P. Gross, and M. Krook, *A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems*, Physical Review, **94**, 511, 1954.
41. A. Narváez, T. Zauner, F. Raischel, R. Hilfer, and J. Harting, *Quantitative analysis of numerical estimates for the permeability of porous media from lattice-Boltzmann simulations*, Journal of Statistical Mechanics: Theory and Experiment, **2010**, P211026, 2010.
42. X. Shan and H. Chen, *Lattice Boltzmann model for simulating flows with multiple phases and components*, Physical Review E, **47**, 1815, 1993.
43. E. Orlandini, M.R. Swift, and J.M. Yeomans, *A lattice Boltzmann model of binary-fluid mixtures*, Europhysics Letters, **32**, 463, 1995.
44. M. R. Swift, E. Orlandini, W. R. Osborn, and J. M. Yeomans, *Lattice-Boltzmann simulations of liquid-gas and binary fluid systems*, Physical Review E, **54**, 5041, 1996.
45. S.V. Lishchuk, C.M. Care, and I. Halliday, *Lattice Boltzmann algorithm for surface tension with greatly reduced microcurrents*, Physical Review E, **67**, 036701, 2003.
46. T. Lee and P.F. Fischer, *Eliminating parasitic currents in the lattice Boltzmann equation method for nonideal gases*, Physical Review E, **74**, 046709, 2006.
47. H. Liu, Q. Kang, C. R. Leonardi, B. Jones, S. Schmieschek, A. Narváez Salazar, J. R. Williams, A. J. Valocchi, and J. Harting, *Multiphase lattice Boltzmann simulations for porous media applications - a review*, Computational Geoscience, **in press**, 2014.
48. C.K. Aidun, Y. Lu, and E.-J. Ding, *Direct analysis of particulate suspensions with inertia using the discrete Boltzmann equation*, Journal of Fluid Mechanics, **373**, 287, 1998.
49. C.K. Aidun and J.R. Clausen, *Lattice-Boltzmann method for complex flows*, Annual Review of Fluid Mechanics, **42**, 439, 2010.
50. A.J.C. Ladd, *Numerical simulations of particulate duspensions via a discretized Boltzmann equation. Part I. Theoretical foundation*, Journal of Fluid Mechanics, **271**, 285, 1994.
51. A.J.C. Ladd, *Numerical simulations of particulate suspensions via a discretized Boltzmann equation. Part II. Numerical results*, Journal of Fluid Mechanics, **271**, 311, 1994.
52. A.J.C. Ladd and R. Verberg, *Lattice-Boltzmann simulations of particle-fluid suspensions*, Journal of Statistical Physics, **104**, 1191, 2001.
53. B.J. Berne and P. Pechukas, *Gaussian model potentials for molecular interactions*, Journal of Chemical Physics, **56**, 4213, 1972.

54. F. Janoschek, F. Toschi, and J. Harting, *Simplified particulate model for coarse-grained hemodynamics simulations*, Physcal Review E, **82**, 056710, 2010.
55. F. Janoschek, J. Harting, and F. Toschi, *Accurate lubrication corrections for spherical and non-spherical particles in discretized fluid simulations*, Submitted for publication, 2014.
56. H. Hertz, *Über die Berührung fester elastischer Körper*, Journal für reine und angewandte Mathematik, **92**, 156, 1881.
57. N. Satofuka and T. Nishioka, *Parallelization of lattice Boltzmann method for incompressible flow computations*, Computational Mechanics, **23**, 164, 1999.
58. J.C. Desplat, I. Pagonabarraga, and P. Bladon, *LUDWIG: A parallel lattice-Boltzmann code for complex fluids*, Computer Physics Communications, **134**, 273, 2001.
59. J. Harting, M. Harvey, J. Chin, M. Venturoli, and P.V. Coveney, *Large-scale lattice Boltzmann simulations of complex fluids: advances through the advent of computational grids*, Philosophical Transactions of the Royal Society of London A, **363**, 1895, 2005.
60. C. Feichtinger, S. Donath, H. Köstler, J. Götz, and U. Rüde, *WaLBerla: HPC software design for computational engineering simulations*, Journal of Computational Science, **2**, 105, 2011.
61. M. Bernaschi, M. Fatica, S. Melchionna, S. Succi, and E. Kaxiras, *A flexible high-performance Lattice Boltzmann GPU code for the simulations of fluid flows in complex geometries*, Concurrency Computat.: Pract. Exper., **22**, 1, 2010.
62. E.T. Coon, M.L. Porter, and Q. Kang, *Taxila LBM: a parallel, modular lattice Boltzmann framework for simulating pore-scale flow in porous media*, Computational Geoscience, **18**, 17, 2014.
63. A. Scheludko, B.V. Toshev, and D.T. Bojadjiev, *Attachment of particles to a liquid surface (capillary theory of flotation)*, Journal of the Chemical Society Faraday Transactions I, **72**, 2815, 1976.
64. A. W. Neumann and J. K. Spelt, *Applied surface thermodynamics*, Marcel Dekker, New York, 1996.
65. P.A. Kralchevsky and K. Nagayama, *Particles at fluid interfaces and membranes*, Elsevier Science, Amsterdam, 2001.
66. A. V Rapacchietta, A. W Neumann, and S. N Omenyi, *Force and free-energy analyses of small particles at fluid interfaces: I. Cylinders*, Journal of Colloid and Interface Science, **59**, 541, 1977.
67. A. V Rapacchietta and A. W Neumann, *Force and free-energy analyses of small particles at fluid interfaces: II. Spheres*, Journal of Colloid and Interface Science, **59**, 555, 1977.
68. P. Singh and D. D. Joseph, *Fluid dynamics of floating particles*, Journal of Fluid Mechanics, **530**, 31–80, 2005.
69. D. D. Joseph, J. Wang, R. Bai, B. H. Yang, and H. H. Hu, *Particle motion in a liquid film rimming the inside of a partially filled rotating cylinder*, Journal of Fluid Mechanics, **496**, 139–163, 2003.
70. I.B. Ivanov, P.A. Kralchevsky, and A.D. Nikolov, *Film and line tension effects on the attachment of particles to an interface: I. Conditions for mechanical equilibrium of fluid and solid particles at a fluid interface*, Journal of Colloid and Interface Science, **112**, 97, 1986.

71. R. Aveyard and J.H. Clint, *Particle wettability and line tension*, Journal of the Chemical Society Faraday Transactions I, **92**, 85, 1996.
72. J. Faraudo and F. Bresme, *Stability of particles adsorbed at liquid/fluid interfaces: Shape effects induced by line tension*, Journal of Chemical Physics, **118**, 6518, 2003.
73. S. Levine, B.D. Bowen, and S.J. Partridge, *Stabilization of emulsions by fine particles I. Partitioning of particles between continuous phase and oil/water interface*, Colloids and Surfaces, **38**, 325, 1989.
74. T. F. Tadros and B. Vincent, *Encyclopedia of emulsion technology, Vol.1, p129*, Marcel Dekker, New York, 1983.
75. J. Guzowski, M. Tasinkevych, and S. Dietrich, *Capillary interactions in Pickering emulsions*, Physical Review E, **84**, 031401, 2011.
76. B. Binks and T. Horozov, *Colloidal Particles at Liquid Interfaces*, Cambridge University Press, Cambridge, 2006.
77. P. G. de Gennes, *Wetting: statics and dynamics*, Reviews of Modern Physics, **57**, 827, 1985.
78. J. F. Joanny and P. G. de Gennes, *A model for contact angle hysteresis*, J. Chem. Phys., **81**, 552, 1984.
79. S. B. G. O'Brien, *The meniscus near a small sphere and its relationship to line pinning of contact lines*, Journal of Colloid and Interface Science, **183**, 51, 1996.
80. G.I. Taylor, *The viscosity of a fluid containing small drops of another fluid*, Proceedings of the Royal Society A, **138**, 41, 1932.
81. G.I. Taylor, *The formation of emulsions in definable fields of flow*, Proceedings of the Royal Society A, **146**, 501, 1934.

Multiparticle Collision Dynamics: Methods and Applications

Marisol Ripoll

Institute of Complex Systems, Research Centre Jülich, 52425 Jülich, Germany

E-mail: m.ripoll@fz-juelich.de

Multiparticle collision dynamics is a particle-based simulation method developed to investigate the properties of mesoscale objects such as colloids or polymer solutions. A large number of works have now consolidated the method, as a very solid alternative for the simulation of most types of soft matter systems. Hydrodynamic interactions, transport of heat and mass, as well thermal fluctuation are by construction included with this technique. In this chapter the details of the method implementation are summarized, and some examples of the simulation of hydrodynamic flows in soft matter systems are reviewed.

1 Introduction

Macroscopic descriptions of fluid properties have been extensively studied on the basis of the work of Navier and Stokes from the first half of the 19th century[71]. The Navier-Stokes equation accounts for the global and local conservation of momentum. The complete description of the fluid requires also the continuity equation for the mass conservation and the heat equation for the energy conservation. In case that thermal fluctuations are of relevance, the three conservation equations can be generalized to the fluctuating hydrodynamic equations[37]. In a few cases these equations can be solved analytically, but in general fluid behavior predictions require to solve coupled sets of nonlinear partial differential field equations by use of finite difference or finite element methods. This is the basis for a well established set of methods which are still widely employed. Nevertheless, there are many situations where these methods can not be applied. This can be due to the no applicability of the continuum equations, to the difficulties in obtaining numerical convergence. Relevant examples of such limitations appear in problems in which the microscopic properties of the fluid components are important in determining the overall fluid behavior.

From a different perspective, simulation methods like molecular dynamics simulation (MD) or Monte Carlo (MC) are based on a microscopic description of fluids[16]. These techniques would reproduce the exact dynamical, equilibrium or steady state properties of the complex fluids if unlimited computing power would be available. To simulate in all detail the dynamical behavior of mesoscale objects, with their surrounding solvent is, due to the large separation of relevant length and time scales, far away from our computational possibilities. However, these intermediate scales (see Fig. 1) are essential to understand a large number of macroscopic phenomena. Most soft matter systems fall into this range of mesoscopic systems ranging from polymers, colloids, droplets, liquid crystals, mixtures, porous media and many biological matter like membranes, or vesicles. Furthermore, many are the processes in which these systems display interesting features like shear flow, microfluidics, microphase separation, or membrane structuring among others. The understanding of these phenomena is still an outstanding challenge for our basic knowledge of fluid mechanics, and moreover, have a considerable practical relevance.

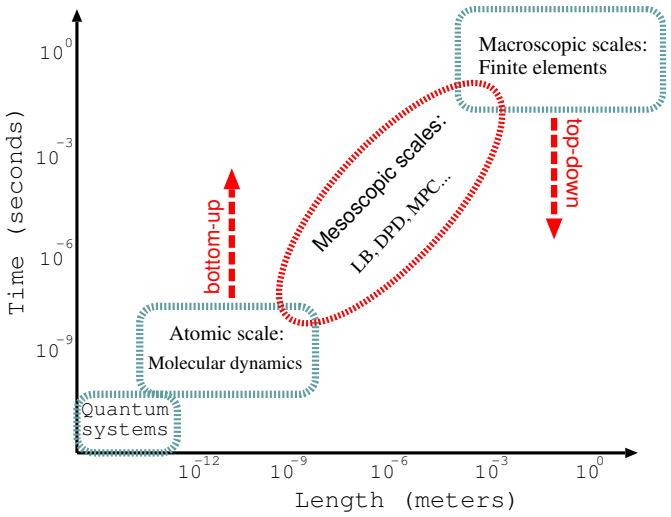
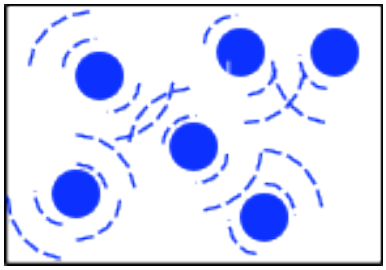


Figure 1: Mesoscale dynamics bridges microscopic and macroscopic scales.

The aim to 'bridge the length and time scales gap', and the increasing availability of computing power, has stimulated the development of several mesoscale simulation techniques in recent years[31,16,4]. This has been done fundamentally either from the 'top-down' approach which consists on discretizations of the continuum equations, or from the 'bottom-up' approach which consist in coarse grained descriptions of the fluid where the microscopic scale is strongly simplified, but relevant effects are still taken into account. In this line, it can be considered that a first mesoscopic model of colloids suspensions is to ignore the solvent and to treat the colloids as hard sphere. The colloids structure is then summarized as impenetrable objects. Their dynamics is accounted by means of MD or MC simulations. This has indeed been successful in predicting static properties of colloidal dispersions in quite early simulations. However, in order to describe dynamical properties of colloidal dispersions, one should also consider solvent effects, and the detail in which this has to be done will depend on the particular effect that one wants to reproduce. One important issue will be if hydrodynamic interactions (HI) should be included or not.

In the absence of *hydrodynamic interactions*, a force $\mathbf{F}_i$ applied to a particle will result in a linear increase of the particle velocity $\mathbf{v}_i = \frac{D_0}{k_B T}\mathbf{F}_i$, where k_B is the Boltzmann constant, T the system temperature, and D_0 the diffusion coefficient[8]. D_0 is a constant that depends on the solute and solvent properties. In the presence of hydrodynamic situations, the velocity increase of a particle due to an applied force is not linear anymore but depends on the position and velocities of all neighboring particles, $\mathbf{v}_i = \frac{1}{k_B T}\sum_j \mathbb{D}_{ij}(\mathbf{r}^N)\mathbf{F}_j$, where $\mathbb{D}_{ij}(\mathbf{r}^N)$ is the hydrodynamic mobility tensor or *Oseen tensor*. Hydrodynamic interactions are therefore dynamical many-body forces characterized by long range interactions that in three dimensions decay like the inverse particle distance, $\sim r^{-1}$. The proper description of such HI is fundamental in many systems and processes like polymer or colloid suspensions, microphase separation or microfluidics devices.

In recent years, several particle-based hydrodynamic mesoscopic techniques have been



When one particle in solution moves, the fluid around is perturbed such that the movement of all neighboring particles is affected.

Figure 2: Diagram of hydrodynamic interactions.

proposed[31]. Mass and momentum local conservation are necessary requirements for hydrodynamic interactions to be correctly modeled. Two of the most well-known and extended hydrodynamic mesoscopic simulation techniques are Lattice-Boltzmann (LB)[46,70] and dissipative dynamics dynamics (DPD)[22,13,20]. In the LB method particle densities move in the nodes of a lattice with discretized velocities, such that mass and momentum fulfil local conservation laws[70]. In standard DPD, particles positions and velocities are updated according to Newton equations of motion by considering conservative, dissipative and stochastic interactions. Mass and momentum are then conserved quantities and thermal fluctuations are taken into account. A more recent approach has been proposed by Malevanents and Kapral[42,43] with the method multiparticle collision dynamics (MPC), also known as stochastic rotation dynamics (SRD). This method conserves in its basic implementation mass, momentum, and energy, including also thermal fluctuations, and the motion is not restricted by the presence of an underlying lattice. This simplicity presents MPC as a promising tool for the investigation of numerous systems.

In this chapter the implementation details of the mesoscopic method MPC are introduced. A couple of example applications are also reviewed which selection is unavoidably biased by my own experience in the field. For further explanations of the method, I refer the interested reader to more extended reviews[89,30,17] and to the original literature.

2 Multiparticle Collision Dynamics: The Method

This mesoscale simulation technique is developed from a previous method known as direct simulation Monte Carlo (DSMC). DSMC was developed for simulating flow of relatively dilute gases[6], and consists in a set of particles that alternate streaming and collision steps. In the streaming step all particles move ballistically, and then collisions are performed among randomly selected pairs of particles. The space is divided into collision boxes inside which the mentioned pairs are chosen. The number of collisions per time step is fixed by the known collision frequency at the specified density. The precise rules for the collisions depend on the molecular model that one wants to reproduce. This is still a widely employed method for the simulation of gases, important for example in aeronautic problems, but mostly restricted to such diluted gases. An extension of the DSMC for dense gases has been recently proposed[1].

Multiparticle collision dynamics is a variant of the DSMC method, in which binary collisions are replaced by multi-particle collisions in a prescribed collision volume. Mass,

momentum, and energy are locally conserved quantities by construction and it has been demonstrated that the hydrodynamic equations are satisfied. MPC can also be consider the most recent mesoscopic technique, since it was first introduced by Malevanets and Kapral[42,43] in 1999. Since there has been an large number of works to consolidate the method, The MPC method has already been extensively tested, showing for example to reproduce the Navier-Stokes equation[42], to include hydrodynamic interactions for example in polymer solutions[64,50], or more recently to fulfil the fluctuation theorem[5].

In this one and a half decades, MPC has successfully been applied to a large number of systems like colloidal solutions[54,79], rod-like colloids[63,62], star polymers[66,69,14], polyelectrolyte solutions[15], red-blood cells[52,47]. These systems have been simulated in various conditions like equilibrium, confinement, shear or capillary flows, or under temperature or density gradients. More recently, MPC is being very successful in the simulations of various systems of microswimmers both biological[12,18,88,61,11] and synthetic[68,7,81,87,59,90]. MPC has furthermore served to propose novel microscale machines[85,80].

2.1 MPC fluid model

The fluid is modeled by N point particles. Each of these particles is characterized by its position $\mathbf{r}_i$ and velocity $\mathbf{v}_i$, and labeled with $i = 1, \ldots, N$. Positions and velocities are continuous variables, which evolve in discrete increments of time. The mass m_i associated with the particles is usually taken to be the same, but more generally, different masses can be assigned.

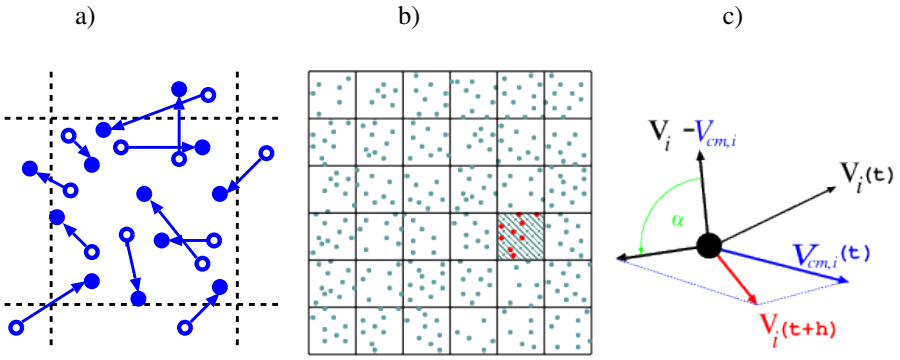


Figure 3: Diagram of the MPC dynamics in 2 dimensions. (a) Streaming step, (b) particles sorted into collision boxes and (c) rotation of the particle velocity relative to the center of mass.

The MPC algorithm consists of two steps, streaming and collision, which are illustrated in Fig. 3. In the streaming step the particles do not interact with each other (see Fig. 3a), they move ballistically according to their velocities during a time increment h, to which I will refer as *collision time*. Thereby, the evolution rule is

$$\mathbf{r}_i(t+h) = \mathbf{r}_i(t) + h\mathbf{v}_i(t). \tag{1}$$

In the collision step, the particles are sorted into collision boxes (see Fig. 3b), and interact with *all* other particles in the same collision box. This multibody interaction takes

place through the collision box center of mass velocity. This is

$$\mathbf{v}_{cm,i}(t) = \frac{\sum_j^{(i,t)} (m_j \mathbf{v}_j)}{\sum_j m_j}, \tag{2}$$

such that $\mathbf{v}_{cm,i}(t)$ is the velocity of the center of mass of all particles j, which are located in the collision box of particle i at the considered time t. The collision boxes are typically the unit cells of a d-dimensional cubic lattice with lattice constant a, although other geometries would be possible. The collision is then defined as a rotation of the velocities of all particles in a box in a co-moving frame with its center of mass. Thus, the velocity of the i-th particle after the collision is

$$\mathbf{v}_i(t+h) = \mathbf{v}_{cm,i}(t) + \mathcal{R}(\alpha) \left[\mathbf{v}_i(t) - \mathbf{v}_{cm,i}(t)\right], \tag{3}$$

where $\mathcal{R}(\alpha)$ is a stochastic rotation matrix. This implies that each particle changes during the collision the magnitude and the direction of its velocity (see Fig. 3c), in such a way that the total momentum and kinetic energy are still conserved within every collision box. This is easy to visualize since the collision box center of mass velocity $\mathbf{v}_{cm,i}$ does not change during the collision,

$$\sum_j^{(i,t)} m\mathbf{v}_j(t+h) = \sum_j^{(i,t)} m \left(\mathbf{v}_{cm,i}(t) + \mathcal{R}(\alpha) \left[\mathbf{v}_j(t) - \mathbf{v}_{cm,i}(t)\right]\right) = \sum_j^{(i,t)} m\mathbf{v}_j(t) \tag{4}$$

$$\sum_j^{(i,t)} \frac{m}{2} v_j^2(t+h) = \sum_j^{(i,t)} \frac{m}{2} \left(\mathbf{v}_{cm,i}(t) + \mathcal{R}(\alpha) \left[\mathbf{v}_j(t) - \mathbf{v}_{cm,i}(t)\right]\right)^2 = \sum_j^{(i,t)} \frac{m}{2} v_j^2(t).$$

Therefore, with the collision rule in Eq. (3), the conservation of mass, local momentum and kinetic energy are guaranteed by construction.

MPC units - In the simulations, N particles are initially placed at random in a cubic system of linear extension L. The average number of particles in a collision box is $\rho = N(a/L)^d$, the scaled number density. Starting from an arbitrary distribution of velocities, only a few steps are required to reach the Maxwell-Boltzmann velocity distribution. The equilibrium temperature T is then given by the average kinetic energy $m\left\langle \mathbf{v}_i^2 \right\rangle = 3k_BT$, where k_B is the Boltzmann constant. It is necessary to choose reference units, and these can for instance be,

Mass unit:	m	$\equiv$ particle mass
Length unit:	a	$\equiv$ collision box size
Energy units:	k_BT	$\equiv$ system temperature

which corresponds to measure length and time according to $\hat{x} = x/a$ and $\hat{t} = t\sqrt{k_BT/ma^2}$. The scaled mean free path is then given by $\lambda = \hat{h}$.

Random rotation - Together with the collision step the stochastic rotation matrix $\mathcal{R}(\alpha)$ has been introduced, such that α is a parameter of the model. The direction around which the rotation is performed is different in neighboring boxes, but inside a particulate box all particles suffer the same rotation. In two dimensions, the rotation of the relative velocity to the box center of mass is simply given by an angle $\pm\alpha$, where the sign independent and randomly chosen for each cell (see Fig. 4a). In three dimensions, two main schemes have

been described for the random collisions. The first one[74] chooses the rotation direction among the three main axis and the rotation is performed by an angle $\pm\alpha$. The second scheme[2] consists in choosing a random direction in space for each box around which the relative velocities are rotated by the angle α. The second scheme is slightly more difficult to implement but in return it introduces less anisotropy in the system due to the underling lattice.

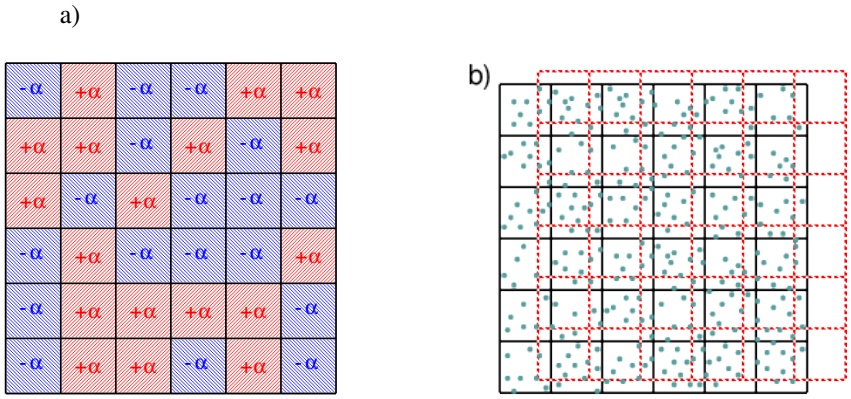


Figure 4: a) Realization of the random direction rotation in two dimensions (red $+\alpha$, blue $-\alpha$). b) Diagram of the random shift of the collision grid.

Random shift - In order to perform the multiparticle collision, as I have already discussed, particles are sorted into cells where the collision takes place. The choice of these collision boxes defines a preferential grid, and the first naive choice would be a fixed grid whose outside borders coincide with the system boundaries, similar as sketched in Fig. 3b. Nevertheless, such a fixed grid does not fulfill Galilean invariance if the mean free path is not large enough. This can be visualized in the case of two systems, one at rest and the second moving with a constant velocity. If the displacement of particle i is smaller than the size of the collision box a, the particles in the collision box will not be the same in the moving system and in the system at rest, which will lead to different dynamics in the two cases and therefore a to breakdown of the Galilean invariance. Therefore, a *random shift* of the collision grid has to be performed in the execution of the collision step[25,26], in order to ensure Galilean invariance in the full range of possible parameters.

Random shift is performed by displacing the collision grid a random number uniformly distributed in the interval $(0, 1)$ which is chosen independently in each collision. In Fig.4b, the solid black grid represents the fixed grid, while the discontinuous red grid would be one of the possible displaced grids. Note that periodic boundary conditions are applied, such that for example, the last-right column of the shifted grid would include some particles of the first-left column of the fixed grid. As a consequence of such a shift no special frame exits and Galilean invariance is restored. Similarly, two particles placed at a quite small distance would not interact if a fix grid would separate them. The random shift implementation produces also that the probability of two particles to interact will be inversely

proportional to their relative distance, in a way similar to a soft range potential. Random shift also facilitates the transfer of momentum between neighboring particles, since the different positions of the grid in two consecutive collision steps make possible transfer of momentum over larger distances per time unit.

Implementation of walls - For many interesting situations, it is necessary to implement of walls as boundary conditions. Two main types of boundary conditions are possible when simulating a solid wall, these are stick and slip. This classification is performed attending the difference of velocity between the fluid and the wall. *Slip boundary conditions* allow that the fluid in contact with the wall and the wall itself have different velocities. *Stick boundary conditions* (also called no-slip) impose a continuous change of velocity between the wall and the fluid. The standard procedure to obtain stick is known as *bounce-back*. When a particle hits the wall its trajectory is reverted and its velocity is inverted (see Fig. 5a). These considerations apply to many simulation techniques.

For simulating fixed walls with MPC, standard bounce-back is applied during the streaming step. This will be enough when the walls exactly coincide with the boundaries of the collision cells. However, the walls will generally not coincide with the cell boundaries, due to special geometries or to random shift. In these cases (see Fig. 5b), the cells in the boundary will be generally partially filled, which will not lead to the desired stick boundary conditions. An efficient solution to overcome this problem has been proposed[36]. The idea is that for all the cells of the channel which are cut by walls and have as a consequence a number of particles n smaller than the average number of particles in the bulk ρ, extra *virtual particles* will be added. The function of these virtual particles is to obtain the behavior of a fluid with constant density and temperature in the cells at the boundaries. The velocities of the virtual-wall particles are drawn from a Maxwell-Boltzmann distribution of zero average velocity and the same temperature T as the fluid. The collision step in Eq. (3) is then carried out with the average velocity of all particles in the cell. Since the sum of random vectors drawn from a Gaussian distribution is again Gaussian-distributed, it is not necessary to calculate the velocities of the individual particles. Instead, the center of mass velocity in Eq.(3) can be written as

$$\mathbf{v}_{cm,i} = \frac{\sum_j^i m\mathbf{v}_j + \mathbf{a}}{\rho} \tag{5}$$

where $\mathbf{a}$ is a vector whose components are numbers from a Maxwell-Boltzmann distribution with zero average and variance $(\rho - n)k_BT$.

Poiseuille flow Giving a fluid resting between to planar walls, a gravitational field g is applied in one direction parallel to the walls. After a relaxation time, the system reaches a stationary state with a parabolic velocity profile between the walls and in the direction of the force. See an example in Fig. 5c, where planar fixes walls are implemented at $y = 0$ and $y = L_y = 25$. It is known[71] that the measured maximum velocity of the parabola is inversely proportional to the kinematic viscosity ν of the fluid like $v_{max} = gL_y^2/(8\nu)$. This behavior is reproduced by MPC simulations, and it can be used as a measurement for the viscosity displayed by the MPC fluid[36,64,65]. Alternative methods to determine the viscosity from simulations have been employed in Refs.[33] and[27], where a system under shear flow and vorticity correlations have been respectively used.

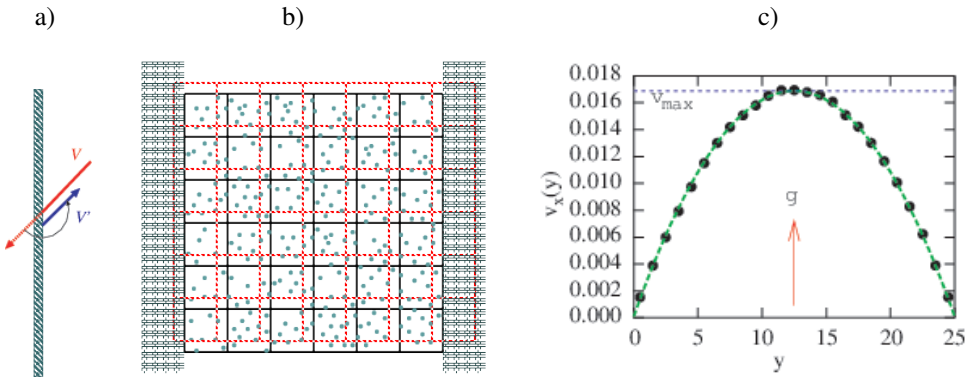


Figure 5: a) Diagram of the bounce-back rule. b) Random-shift in the presence of walls. c) Parabolic velocity profile in the presence of a gravitational field reproduced by MPC.

2.2 Transport properties

Transport coefficients of the MPC solvent have been studied intensively. Analytical expressions have been derived from kinetic theory by generalizing point-like collisions to finite collision volumes[26,27,74,33]. The theoretical expressions describe numerical results very well. This is quite convenient since different properties of the fluid can be tuned by choosing the adequate set of model parameters.

Viscosity - The total kinematic viscosity, $\nu = \nu_{kin} + \nu_{coll}$, is the sum of two contributions, the kinetic viscosity ν_{kin} and the collisional viscosity ν_{coll}, which have been calculated in two and three dimensions. In three dimensions[74,33], these expressions are

$$\nu_{coll} = \frac{1}{\lambda}\frac{(1-\cos\alpha)}{18}\left(1-\frac{1}{\rho}\right) \tag{6}$$

$$\nu_{kin} = \lambda\left[\frac{1}{(4-2\cos\alpha-2\cos 2\alpha)}\frac{5\rho}{\rho-1}-\frac{1}{2}\right].$$

The viscosity data obtained from Poiuselle flow simulations are presented in Fig. 6 together with the theoretical predictions of Eq. (6). The obtained agreement is quite remarkable, in contrast to the case of other mesoscopic simulation techniques such as dissipative particle dynamics[57]. Density fluctuations can also be included in the theory[33], which noticeably improves the agreement with the simulations results for small number densities; for $\rho \geq 5$ these contributions are negligible.

The ratio between the kinetic and the collisional contributions to the kinematic viscosity varies considerably with the model parameters, as can be seen easily from the theoretical expressions (6). In Fig. 6 the total kinematic viscosity and its two contributions are plotted as a function of the collision time step. The collisional contribution is dominant for large collision small collision times, while the kinetic viscosity dominates in the opposite case of small collision angles and large collision times.

Diffusion coefficient - The self-diffusion coefficient can be a calculated by the Green-Kubo formalism from the velocity autocorrelation function (VACF) as

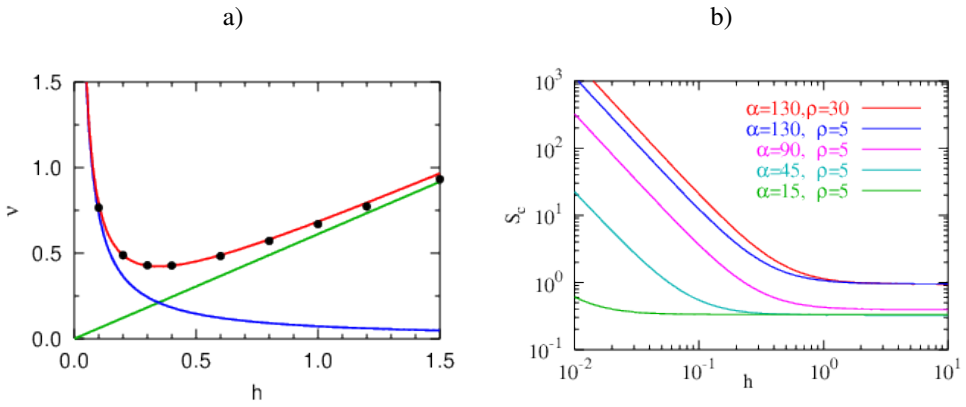


Figure 6: a) Dimensionless kinematic viscosity as a function of the collision time λ. Red line is the total theoretical prediction with $\alpha = 130$ and $\rho = 5$, blue line is the collisional contribution and green line the kinetic contribution. Symbols are the simulation results, with system size $L/a = 20$. B) Theoretical Schmidt number versus collision time. The α and ρ parameters are specified in the legend.

$D = \frac{1}{3}\int_0^\infty dt\, \langle \mathbf{v}(t)\mathbf{v}(0)\rangle$[27]. In order to obtain an analytical prediction, the Brownian approximation is employed for the VACF, what yields to,

$$D = \lambda\left(\frac{1}{\gamma} - \frac{1}{2}\right), \qquad \gamma = \frac{2}{3}(1-\cos\alpha)\left(1 - \frac{1}{\rho}\right), \tag{7}$$

where the decorrelation factor γ has been specified for the 3-dimensional case. A simulation measurement can be obtained through the VACF or by calculating the averaged mean square displacement. The agreement with simulation results is quite reasonable for large values of the mean free path λ, but not so for small ones. This is due to the fact that the Brownian approximation neglects the effect of correlations between particles. This deviation is therefore an indication of the importance of the enhancement of the collective behavior in the regime where the mean free path is smaller than the box size[64,65].

Schmidt number - A convenient measure of the importance of hydrodynamics is the Schmidt number $Sc = \nu/D$, where ν is the kinematic viscosity and D the diffusion coefficient. Thus, Sc is the ratio between momentum transport and mass transport. It is known that this number for gases is on the order of unity, while in fluids like water it is on the order of 10^2 to 10^3. A prediction for the Schmidt number of a MPC fluid can be obtained from the theoretical expressions (6) for the kinematic viscosity, and the diffusion coefficient in Eq. (7). In Fig. 6, the theoretical prediction is plotted for Sc as a function of the mean free path (or collision time) for different values of the rotation angle. This shows that Sc becomes considerably larger than unity for the same range of parameters where the collisional viscosity is considerably larger than the kinetic viscosity.

In order to perform simulations where the hydrodynamic effects are more easily taken into account, it is advisable to employ large values of the rotation angle α (like $\alpha = 130^o$) and small values of the collision time (like $h = 0.1$ or smaller). These restrictions, together with the election of the number density ρ, still give a large margin to chose particular values for the transport coefficients like the viscosity $\eta = \nu\rho$. However, it should be recall that

the simulations will become computationally more expensive as for smaller collision times and for larger number densities.

2.3 Model variations

MPC with Andersen thermostat
The MPC algorithm has been constructed such that mass, and translational momentum are locally conserved, which is essential for obtaining correct hydrodynamic behavior. As already discussed, the original version of MPC defined with the collision rule in Eq. (3) also conserves energy. This means that the method is suited to be applied to investigate transport of energy or systems with temperature gradients[40], and in principle also that temperature is constant without the necessity of any additional thermostat. This is though not the case only in many cases the system has external influences like in the presence of a gravitational force, shear flow, or an electric field. In these cases, the interaction with a thermostats might be necessary similar to most simulation techniques[23,24]. A collision with an intrinsic thermostat can be defined[53] alternatively to Eq. (3) as

$$\mathbf{v}_i(t+h) = \mathbf{v}_{cm,i}(t) + \delta\mathbf{v}^{ran} = \mathbf{v}_{cm,i}(t) + \mathbf{v}_i^{ran} - \frac{1}{N_c}\sum \mathbf{v}_j^{ran} \tag{8}$$

where N_c is the number of particles in the collision cell and the sum runs over all particles in the cell. Here, the new relative velocities are Gaussian random numbers with variance $\sqrt{k_B T/m}$, being T the thermostating temperature. This method is know as MPC-AT in contrast to the standard one which is now frequently called MPC-SRD.

MPC with angular momentum conservation
Besides the multiple advantages of the two described MPC methods, local angular momentum is not conserved in any of them. Although the violation of angular momentum conservation does not affect the generic form of the hydrodynamic equations[28,23,55], the fluid stress tensor is no longer symmetric[28,60,19,86]. This affects the propagation of sound, but is unimportant for the evolution of vorticity of a bulk fluid. The absence of angular momentum conservation leads to a non-physical contribution to the torque in the case of circular Couette flow, or to incorrect velocity fields in some cases[19,86].

Angular momentum conservation can be restored by imposing constraints on the new relative velocities. The collision rule in this case is

$$\mathbf{v}_i(t+h) = \mathbf{v}_{cm,i}(t) + \mathbf{v}_i^{ran} - \frac{1}{N_c}\sum \mathbf{v}_j^{ran} + \left\{\frac{m}{\Pi}\sum \left[\mathbf{r}_{j,c} \times \left(\mathbf{v}_j - \mathbf{v}_j^{ran}\right)\right] \times \mathbf{r}_{i,c}\right\}, \tag{9}$$

where Π is the moment of inertia tensor of the particles in the cell, and $\mathbf{r}_{j,c} = \mathbf{r}_j - \mathbf{R}_c$ is the relative position of particle j in the cell and $\mathbf{R}_c$ is the center of mass of all particles in the cell. This is then an angular-momentum conserving modification of MPC-AT[53,19], denoted MPC-AT+a.

Non-ideal MPC
A new version of the model proposed in Ref.[29] introduces generalized excluded volume interactions between fluid particles such that fluids with non-ideal-gas equation of state are produced. The additional consideration of interactions between different types of particles leads to the possibility of simulating multicomponent fluids[73].

2.4 Implementation of complex structures

After regarding the implementation and characteristic behavior of a simple fluid with simulated with MPC, the next important question is how to model complex structures with a MPC solvent. The general strategy is to define an hybrid algorithm where the solvent is simulated with the MPC technique and both, the solute description and the solute-solvent interactions, are accounted with specific models. Different structures like colloids, polymers or membranes have therefore developed different coupling schemes. Some of them are summarized below.

MD coupling This was the first coupling introduced for MPC by Malevanets and Kapral[43,44], it was applied to a colloidal dispersion. The idea of this coupling is that both the solute-solute and solute-solvent interactions are taken into account through explicit potentials with molecular dynamics (MD), and only the solvent-solvent interactions were mesoscopically described through MPC. This leads to a considerable speed up of the simulations respect to a simulation that will consider all the solvent-solvent interaction with MD.

For example, in case a polymer is considered, N_m particles of mass M are taken to be the polymer monomers. These monomers are connected by harmonic springs and have excluded-volume interactions with the solvent particles and eventually also between monomers of the same polymer. The MD time steps are integrated, for instance, with the velocity-Verlet algorithm with a time step Δt generally 20 to 100 times smaller than the collision time. This discrete time evolution is then interrupted every $h/\Delta t$ steps such that solvent particles interchange momentum among themselves through a MPC collision (see Fig. 7a). Typical potential parameters are then related to the MPC units.

In Ref.[38] Lee and Kapral employ this coupling to study dynamics of polymer with solvent of different qualities. This is achieved by changing the excluded volume interactions from purely repulsive to a Lennard-Jones interaction that accounts for repulsive interactions at short distances to attractive interactions at intermediate distances. The poor solvent is reproduced by considering attractive interactions between monomers and repulsive between solvent and solute particles. Oppositely, the good solvent considers repulsive interactions between monomers and repulsive between monomers and solvent particles.

MPC coupling In this case only solute-solute interactions are described with MD, while both solvent-solvent and solute-solvent interactions are described with MPC interactions. This means that the MPC streaming step in Eq. (1) is used to update the positions just of the solvent particles, while in the collision step in Eq. (3) both solvent and solute particles are taken into account. The collision is performed through the collision box center of mass velocity Eq. (2) where the solute particles are generally considered to have larger mass than the surrounding solvent particles. Similarly to the MD coupling, the position update of the solute particles is performed in several MD steps between MPC collisions, and it is in the MD where the particular characteristics of the interaction potential are taken into account.

Coming back to the polymer example[45,50], N_m point particles of mass M are also taken to be the polymer monomers. These monomers are connected by harmonic springs and may also interact via an excluded-volume potential. The typical distance between consecutive monomers (bond length) could be equal to the box size, so that there is no

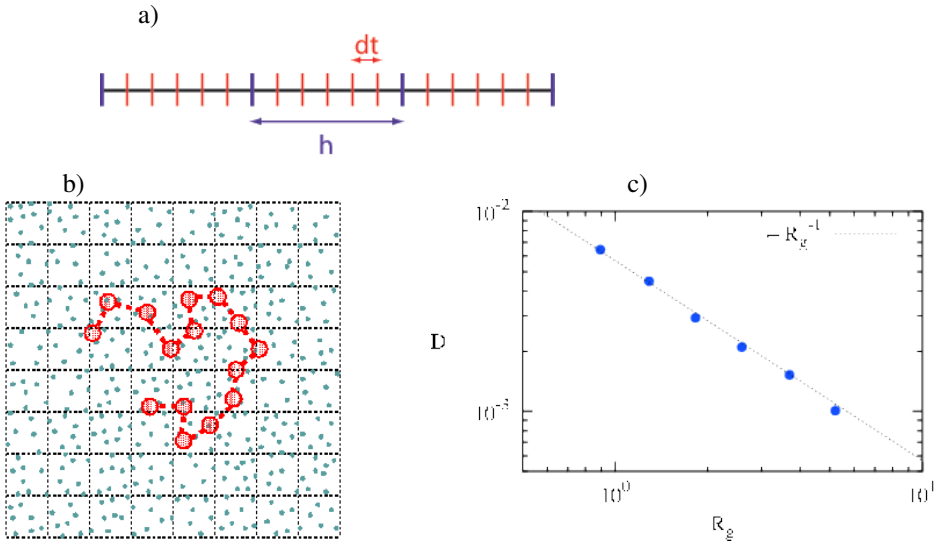


Figure 7: a) Diagram of the time evolution in MD and MPC couplings. b) Diagram of the inclusion of a polymer in a MPC solvent. c) Polymer center of mass diffusion coefficient with the MPC coupling together with the Zimm theoretical prediction.

more than one monomer in each collision box. The potential strength would similarly be taken in units of the thermal energy k_BT. This procedure is interrupted every $h/\Delta t$ steps for updating the positions of the fluid particles and for the MPC collision, where solvent and solute particles interchange momentum. This implies that the solvent particles can enter the cores of the monomers, but the monomers cannot inter penetrate each other (see Fig. 7b). In return, the MPC coupling is computationally less costly than the MD coupling since in every MD time step the solvent is ignored. Other properties like solvent induced depletion interactions are different between both couplings.

The MPC coupling has been used to study the dynamics of short polymer chains in solution[64,50]. The predictions of the Zimm theory, in which the effect of hydrodynamic interactions is taken into account, are in excellent agreement with the simulation results for instance, for the center of mass diffusion coefficient (see Fig. 7c) and for the relaxation times of the Rouse modes. This agreement is found in the parameter regime where the collective behavior has been described to dominate the dynamics, which also is when the Schmidt number displayed larger values. Polymer collapse has also been studied with this coupling showing the relevance of hydrodynamics interactions[32,34].

Boundary conditions coupling Specific solvent-solute interactions can be defined at the boundaries between the solute surfaces and the surrounding solvent. In order to simulate the interaction of a flow with an spherical obstacle, the MPC solvent particles interact with the sphere surface as with a planar wall with a bounce back collision[36,35,2]. As can be seen in Fig. 8 the fluid instability expected for high Reynolds numbers is properly reproduced.

More elaborated rules have been introduced when simulating moving colloids[56]. These interactions take into account not only the transfer of linear momentum and energy be-

tween solvent and solute particles, but also the transmission of angular momentum from the solvent collisions to the solute particle.

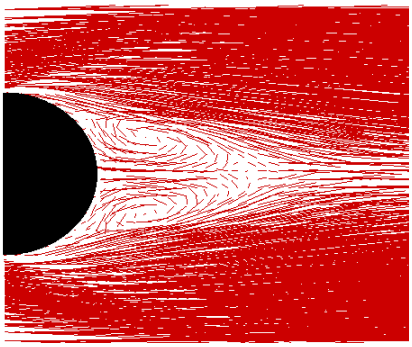


Figure 8: Velocity field of a left-to-right flow past a sphere for high Reynolds number conditions (reproduced from Ref.[2]).

Further couplings In order to describe different specific situations other couplings can be defined. An example is the simulation of vesicles in flow[51,52]. In this case, solvent free models for the simulation of membranes like discretized curvature models or triangulated surfaces are employed in combination of the MPC solvent. Interestingly in this case, the fluids inside and outside the vesicle do not mix and can have different fluid properties like density or viscosity.

2.5 Random MPC solvent

A standard procedure to determine the relevance of HI is to compare simulations with and without HI. In principle existing methods without HI, like BD in Sect. 2.1, would provide a valuable comparison. Nevertheless, the ideal to perform a good comparison is that the considered methods differ as little as possible apart from the inclusion of HI. With this spirit we proposed the *random MPC* solvent[67]. The idea is to replace the solvent-solute interaction in the collision step by the interaction with a heat bath. Since the positions of the solvent particles within a cell are not required in the collision step, no explicit particles have to be considered. Instead, each monomer is coupled with an effective solvent momentum $\mathbf{P}$ which is directly chosen from a Maxwell-Boltzmann distribution of variance $m\rho k_B T$ and a mean given by the average momentum of the fluid field – which is zero at rest, or $(m\rho\dot{\gamma}r_y^i, 0, 0)$ in the case of an imposed shear flow. This strategy is related to that proposed in Sect. 3.1 and Ref.[36] in order to model the no-slip boundary conditions of solvent particles at a planar wall. The total center-of-mass velocity is then given by

$$\mathbf{v}_{cm,i} = \frac{M\mathbf{v}_i + \mathbf{P}}{m\rho + M}, \tag{10}$$

with which the collision step in eq. (3) is performed. Note that similarly to the MPC solvent, the complex structure trajectory is sampled by MD and the interaction with the solvent is performed every collision time h.

The random MPC solvent has therefore similar properties as the MPC solvent, except those associated with HI. The relevant parameters in both methods are the density ρ, the rotation angle α, and the collision time h which can be chosen to be the same. Simple to compare are the velocity autocorrelation functions[65]. The random MPC solvent will only lead to an exponentially decaying correlation function. Naturally, this leads to differences in the diffusion coefficient. Other transport coefficients such as the viscosity depend on HI only weakly[72] and consequently are expected to be essentially identical in both solvents. A very important advantage provided by this method is that *no* explicit solvent particles need to be considered. This translates into a computational speed-up by several orders of magnitude, depending on the considered system size. Moreover, to implement non-equilibrium conditions is straightforward. Examples of application of this technique can be found in Refs.[67,10,39].

3 Simulation of Hydrodynamic Flows with MPC

The described MPC method simulates an explicit solvent which includes hydrodynamic interactions. These interactions have shown to be the consequence of the momentum conservation in the MPC dynamics, where no specific fluid behavior is imposed. The flow is hydrodynamic and results therefore of the specific boundary conditions of the problem. In the following, a few examples of these flow are discussed. For comparison, analytical expressions for hydrodynamic flows can obtained as solution of the Navier-Stokes equations with adequate boundary conditions. Systems of interest are frequently on the steady state systems, with small Reynolds and Mach numbers, such that time derivatives and non-linear terms can be neglected and the Stokes equation applies.

3.1 Colloid exposed to an external force

We first consider the case of a colloid with slip boundary conditions experiencing a constant external force and moving at velocity u. A relevant example is a sedimenting colloidal sphere in a gravitational field. The flow field has to satisfy the three boundary conditions: (i) Its normal component vanishes at the particle surface in the co-moving reference frame. (ii) The slip boundary implies that the tangential stress vanishes at the particle surface. (iii) The flow field vanishes at an infinite distance. The solution of this flow depends then on the shape of the stress tensor and therefore on the presence or absence of angular momentum conservation. Figure 9a displays the velocity field around the colloid. Fig. 9b shows the quantitative comparison of the measured flow with the corresponding analytical expressions in the absence or presence of AMC[86]. The agreement between the analytical expressions and the simulation results is very good.

3.2 Thermophoretic flows

In the presence of a temperature gradient, a colloid in solution shows a directed motion, this is known as thermophoresis[76,58,78]. The colloid motion is not the result of an externally applied force, but it results from the unbalance of the force exerted by the surrounding fluid[83]. In the MPC simulation the thermophoretic force is not externally imposed either,

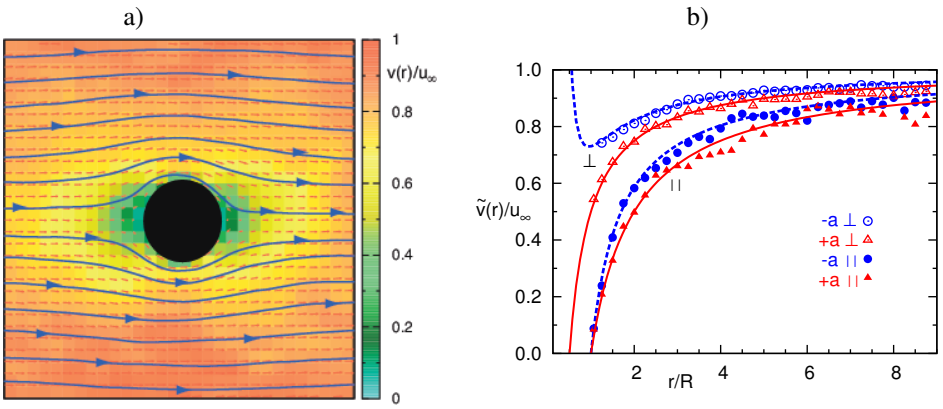


Figure 9: a) Velocity field of an externally applied fluid flow past a fixed colloidal sphere (cross-section through the colloid center). Blue solid lines correspond to the stream lines, small red arrows are the solvent velocity field, and the background color code indicates the normalized velocity modulus. b) Normalized fluid velocity as a function of the separation from the center of a fixed colloidal particle in a fluid with a constant flow velocity at infinity. Symbols refer to the simulation results, in an axis along the flow (||, solid), and axis perpendicular to the flow (⊥, empty). Lines correspond to the analytical predictions.

but it emerges as a consequence of the temperature variation of the solvent colloid interactions.

The temperature gradient can be imposed to the MPC fluid at boundary walls or with periodic boundary conditions, similar as it is done with MD simulations[21,49]. Since MPC-SRD conserves by construction local kinetic energy, temperature inhomogeneities can be simply taken into account, in contrast to other intrinsically isothermal methods as dissipative particle dynamics or lattice Boltzmann, where generalized version of the methods are necessary. The fact that MPC has the equation of state of an ideal gas implies that the temperature gradient is accompanied of a density gradient[40]. Compressibility effects will then be present although they have show though to be of small importance for thermophoretic colloids[84]. The colloid is implemented with MD coupling with the MPC solvent. Due to their tunable properties and simplicity, Lennard-Jones type of potentials have been employed[41] for the colloid solvent interactions. Thermophobic colloids are those that preferably drift to colder areas, as it is the case in most of the experimental colloidal particles. Colloids simulated with attractive colloid solvent interactions and MPC have shown to be thermophobic. In contrast, colloids simulated with repulsive interactions result to be thermophilic, drifting to warmer areas.

Besides the colloid drift, the temperature inhomogeneity also induces a flow of the solvent around the colloid. This flow directly emerges from the MPC simulations due to its hydrodynamic character. The analytical expression of the flow around a fixed colloid can be now obtained from the Stokes equation by considering that the integral of the stress tensor over the particle surface corresponding to the thermophoretic force $\mathbf{f}_T$. The resulting stationary flow field[9,48] is,

$$\mathbf{v}(\mathbf{r}) = -\frac{1}{8\pi\eta r}(\hat{\mathbf{r}}\hat{\mathbf{r}} + \mathbf{I}) \cdot \mathbf{f}_T + \frac{R^2}{8\pi\eta r^3}(3\hat{\mathbf{r}}\hat{\mathbf{r}} - \mathbf{I}) \cdot \mathbf{f}_T. \tag{11}$$

Here, R is the particle radius, η the solvent dynamic viscosity, $\hat{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$ and $\mathbf{I}$ the unit tensor. The flow is the superposition of a Stokeslet and a source-dipole. Equation (11) indicates that the flow velocity around a fixed particle in a temperature gradient has opposite direction to the thermophoretic force, and that it is of long range since it decays linearly with the inverse distance from the particle center. The flow field induced by a thermophilic particle fixed between walls at different temperatures obtained by simulations is shown in Fig. 10. A very precise quantitative agreement with Eqs. (11) has been demonstrated[84] when the effect of back flow is considered, as can be seen Fig. 10. Note that the difference between the cold and the warm branches is small what implies that the compressibility effects in MPC are small.

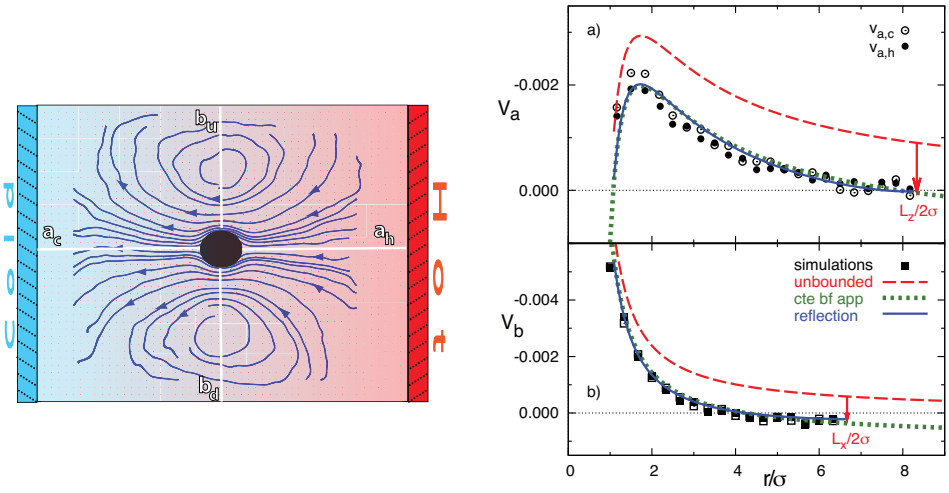


Figure 10: Left) Cross section of the flow field induced by a thermophilic colloid fixed between parallel cold and hot walls. Small red arrows indicate the flow velocity direction and intensity, while the thick blue lines correspond to the flow stream lines. The axes where the flow velocities are quantified in are displayed here in white. Right) Flow velocity as a function of distance from the colloid center. Symbols correspond to simulation results, dashed lines to the theoretical calculation in Eq. (11), dotted lines to the constant backflow approximation, and solid lines to theoretical calculation with the reflection method. Arrows indicate the position of the system boundaries. a) Velocity in axis a. b) Velocity in axis b.

In case the colloid can freely move, the thermophoretic force induces a drift in the colloid. The colloid thermophoretic velocity $\mathbf{u}_T$ results from the balance of the thermophoretic and the friction forces $\mathbf{f}_T = \gamma \mathbf{u}_T$, with γ the friction coefficient, which can be approximated to be spatially constant[82]. The solution of Stokes equation needs to consider now the particle drift and that the balancing forces on the particle result in a vanishing integral of stress tensor over the particle surface. The obtained velocity flow field[3,77] reads,

$$\mathbf{v}(\mathbf{r}) = \frac{R^3}{2r^3}(3\hat{\mathbf{r}}\hat{\mathbf{r}} - \mathbf{I}) \cdot \mathbf{u}_T. \tag{12}$$

The flow velocity across the colloidal center and along the temperature gradient has now the same direction as the thermophoretic force, and decays with the inverse of the distance

cubed. This is much faster than in the case of a fixed particle. The flow field induced by a freely moving colloid in a temperature gradient has been obtained by simulations and it is shown in Fig. 11. A very precise quantitative agreement with Eq. (12) has been found[84] and it is shown in Fig. 11.

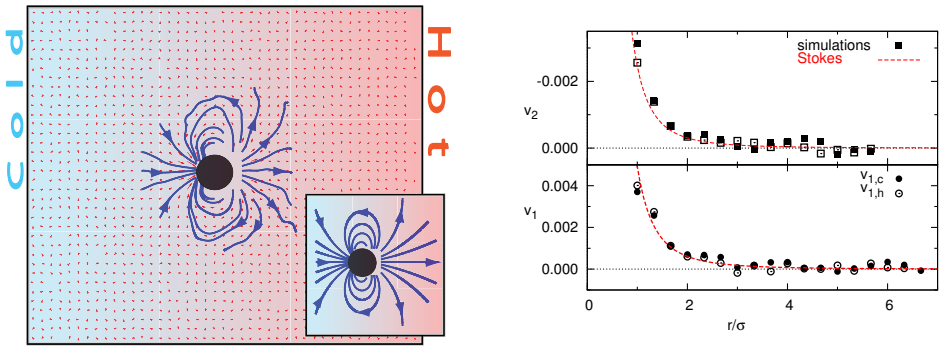


Figure 11: Left) Cross section of the velocity field around a freely drifting thermophilic colloidal sphere in a temperature gradient. The inset shows the theoretical result from Eq. (12). Right) Velocity field around a drifting thermophoretic particle as a function of distance from the particle, with positive direction toward the hot side. Symbols refer to the simulation results, lines to the theoretical calculation from Eq. (12). a) Velocity field V_a, solid and open symbols correspond to V_{ac} and V_{ah}, respectively. b) Velocity field V_b.

Furthermore, the existence of a thermophoretically induced fluid flow has interest not only from a fundamental point of view, but can also find numerous practical applications, like the existence of inter-colloidal hydrodynamic attraction, and the possibility of designing thermophoretic pumps[84,85].

3.3 Thermophoretic crystals

When freely moving colloidal particles are confined between walls at different temperatures, the directional thermophoretic force will naturally drive them towards one of the walls. The colloids may then stay at an averagely fixed distance of the confining wall performing a two-dimensional Brownian motion[75,9]. It is then to be expected that the thermophoretically induced flow field is related to that of a fixed particle in Eq.(11), but asymmetrically modified by the presence of the wall. In contrast to the symmetric system in Fig. 10 where the colloid was fixed equidistant from both walls, the thermophoretically induced flow field of a thermophilic colloid close to the hot wall in Fig. 12 has now a strong lateral component toward the colloidal sphere and parallel to the wall. If a second particle is in the neighborhood of the colloid, it will suffer a hydrodynamic drag toward the first particle, this is a thermophoretically induced attraction force. This attraction force is then perpendicular to the temperature gradient, and will decrease as function of the distance. When the lateral flow is strong enough (high $|\alpha_T|$), the attraction force can be larger than other repulsive contributions or than thermal fluctuations, originating stable colloid aggregation. Such 2D colloidal thermophoretic crystals, have indeed been experimentally observed[75,9], and show to form and be stable only in the presence of a temperature gradient.

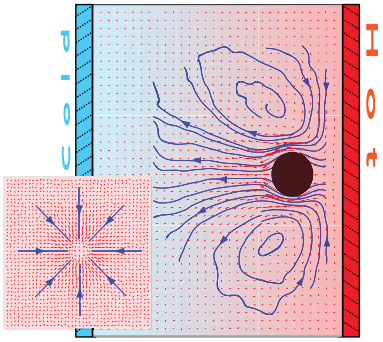


Figure 12: Cross section of the flow field around a thermophilic particle fixed in a near-wall environment. The inset shows the velocity field in a plane parallel to the walls and between the colloid and the hot wall.

3.4 Star polymers under shear

One example in which it is not easy to provide an analytical prediction of the hydrodynamic flow is the interaction of a star polymer with a shear flow. The distortion of the external flow has been investigated with MPC simulations[66,67] and it is shown in Fig. 13a. The shear flow produces the elongation, orientation and rotation of the star polymer with the flow. In correspondence, the fluid shows to rotate together with the star polymer. Outside the region covered by the star polymer, the fluid adapts to the central rotation by generating a counter-rotating vortex and correspondingly two hyperbolic stagnation points of vanishing fluid velocity. It can be understood this fluid behavior will have an effect in the star polymer dynamics, in comparison with a polymer that would be simulated in an unperturbed linear velocity profile expectable from shear flow. The precise effect of the hydrodynamic interactions can in this case be quantified by performing additional simulations in the absence of HI as described in Sec. 2.5, as shown in Fig. 13b for the star polymer rotational frequency. The rotation frequency ω_z and the various functionalities

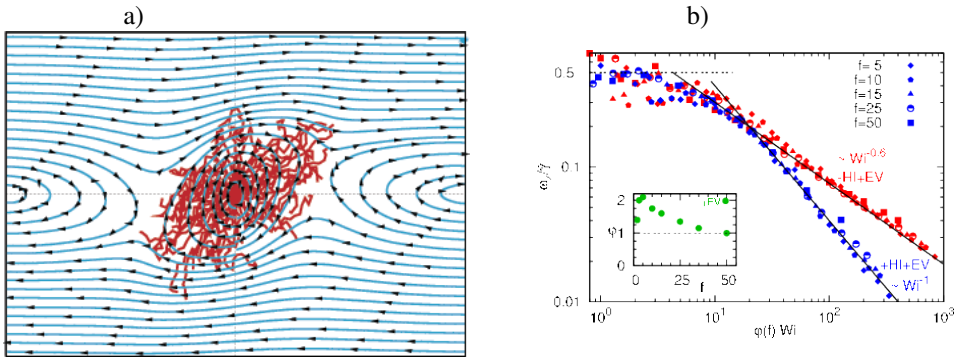


Figure 13: a) Stream lines of a fluid interacting with a star polymer under a shear flow. b) Normalized rotational frequency of star polymers as a function of the applied shear flow in the presence and absence of hydrodynamic interactions.

collapse on two master curves, when the Weissenberg number Wi is rescaled by a factor $\varphi(f)$, which turns out to be the same with and without HI. For small Wi, all frequencies are expected to approach $\omega_z = 0.5\dot{\gamma}$. For large Wi, the master curves follow a power law decay, $\omega_z/\dot{\gamma} \sim Wi^{-\theta}$, with $\theta = 1$ in the presence of HI, and $\theta = 0.6$ without HI. This implies that without HI the rotation frequency keeps growing with the applied shear rate as $\omega_z \sim \dot{\gamma}^{0.4}$, while with HI it approaches a constant for large shear rates. In both cases, the effect of the deformation and alignment under shear flow leads to a sub-linear growth with Wi. In the presence of HI, the co-rotating fluid core, dragged along by the star polymer, screens the driving force even more and a limiting velocity appears, similar to the motion of a capsule. The rotational dynamics is therefore not only quantitatively different in the cases with and without HI, but also qualitatively.

4 Concluding Remarks

In the last fifteen years, multiparticle collision dynamics has shown to be a very convenient and robust simulation method to investigate the dynamical properties of soft matter and biological systems. The simplicity of its implementation, and its versatility is pushing the method forward for a large class of systems where thermal fluctuations are of importance. One additional advantage is that not only the hydrodynamic effects on the solute properties can be studied, but also the distortion that the fluid undergoes due to the presence of the solute.

The implementation details of MPC are comprehensibly summarized in this chapter. A few relevant examples of applications of MPC to systems where the induced flow fields are of relevance are here also reviewed. In the cases where comparison with the solution of the Stokes equation is possible, MPC simulation results show to perfectly agree. Consequently, in cases where analytical solution can not be obtained the results obtained with MPC simulations are very reliable.

References

1. F. J. Alexander, A. L. Garcia, and B. J. Alder. A consistent Boltzmann algorithm. *Phys. Rev. Lett.*, 74:5212–5215, 1995.
2. E. Allahyarov and G. Gompper. Mesoscopic solvent simulations: Multiparticle-collision dynamics of three-dimensional flows. *Phys. Rev. E*, 66:036702, 2002.
3. J. L. Anderson. Colloid transport by interfacial forces. *Annu. Rev. Fluid Mech.*, 21:61, 1989.
4. N. Attig, K. Binder, H. Grubmüller, and K. Kremer, editors. *Computational Soft Matter: From Synthetic Polymer to Proteins*. NIC-Directors, 2004.
5. M. Belushkin, R. Livi, and G. Foffi. Hydrodynamics and the fluctuation theorem. *Phys. Rev. Lett.*, 106:210601, 2011.
6. G. A. Bird. *Molecular Gas Dynamics*. Clarendon, Oxford, 1976.
7. P. de Buyl and R. Kapral. Phoretic self-propulsion: a mesoscopic description of reaction dynamics that powers motion. *Nanoscale*, 5:1337, 2013.
8. J. K. G. Dhont. *An Introduction to Dynamics of Colloids*. Elsevier, 1996.
9. R. Di Leonardo, F. Ianni, and G. Ruocco. Colloidal attraction induced by a temperature gradient. *Langmuir*, 25:4247, 2009.

10. J. Elgeti and G. Gompper. Self-propelled rods near surfaces. *Europhys. Lett.*, 85:38002, 2009.
11. J. Elgeti and G. Gompper. Emergence of metachronal waves in cilia arrays. *Proc. Natl. Acad. Sci.*, 110:4470–4475, 2013.
12. J. Elgeti, U. B. Kaupp, and G. Gompper. Hydrodynamics of sperm cells near surfaces. *Biophys. J.*, 99:1018, 2010.
13. P. Español and P. Warren. Statistical mechanics of dissipative particle dynamics. *Europhys. Lett.*, 30:191, 1995.
14. D. A. Fedosov, S. P. Singh, A. Chatterji, R. G. Winkler, and G. Gompper. Semidilute solutions of ultra-soft colloids under shear flow. *Soft Matter*, 5:4109, 2012.
15. S. Frank and R. G. Winkler. Polyelectrolyte electrophoresis: Field effects and hydrodynamic interactions. *Europhys. Lett.*, 83:38004, 2008.
16. D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press, San Diego, 2nd edition, 2002.
17. G. Gompper, T. Ihle, D. M. Kroll, and R. G. Winkler. Multi-particle collision dynamics: A particle-based mesoscale simulation approach to the hydrodynamics of complex fluids. *Adv. Polym. Sci.*, 221:1–87, 2009.
18. I. O. Götze and G. Gompper. Mesoscale simulations of hydrodynamic squirmer interactions. *Phys. Rev. E*, 82:041921, 2010.
19. I. O. Götze, H. Noguchi, and G. Gompper. Relevance of angular momentum conservation in mesoscale hydrodynamics simulations. *Phys. Rev. E*, 76:046705, 2007.
20. R. D. Groot and P. B. Warren. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.*, 107:4423–4435, 1997.
21. B. Hafskjold, T. Ikeshoji, and S. K. Ratkje. On the molecular mechanism of thermal diffusion in liquids. *Molecular Phys.*, 80:1389–1412, 1993.
22. P. J. Hoogerbrugge and J. M. V. A. Koelman. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhys. Lett.*, 19:155–160, 1992.
23. C. Huang, A. Chatterji, G. Sutmann, G. Gompper, and R. Winkler. Cell-level canonical sampling by velocity scaling for multiparticle collision dynamics simulations. *J. Comp. Phys.*, 229:168, 2010.
24. C. Huang, A. Varghese, G. Gompper, and R. Winkler. Thermostat for nonequilibrium multiparticle-collision-dynamics simulations. *Phys. Rev. E*, 91:013310, 2015.
25. T. Ihle and D. M. Kroll. Stochastic rotation dynamics: A galilean-invariant mesoscopic model for fluid flow. *Phys. Rev. E*, 63:020201(R), 2001.
26. T. Ihle and D. M. Kroll. Stochastic rotation dynamics I: Formalism, galilean invariance, green-kubo relations. *Phys. Rev. E*, 67:066705, 2003.
27. T. Ihle and D. M. Kroll. Stochastic rotation dynamics II: Transport coefficients, numerics, long time tails. *Phys. Rev. E*, 67:066706, 2003.
28. T. Ihle, E. Tüzel, and D. M. Kroll. Equilibrium calculation of transport coefficients for a fluid-particle model. *Phys. Rev. E*, 72:046707, 2005.
29. T. Ihle, E. Tüzel, and D. M. Kroll. Consistent particle-based algorithm with a non-ideal equation of state. *Europhys. Lett.*, 76:664–670, 2006.
30. R. Kapral. Multiparticle collision dynamics: Simulation of complex systems on mesoscales. *Adv. Chem. Phys.*, 140:89–146, 2008.
31. M. Karttunen, I. Vattulainen, and A. Lukkarinen, editors. *Novel Methods in Soft Matter Simulations*. Springer, 2004.

32. N. Kikuchi, A. Gent, and J. M. Yeomans. Polymer collapse in the presence of hydrodynamic interactions. *Eur. Phys. J. E*, 9:63–66, 2002.
33. N. Kikuchi, C. M. Pooley, J. F. Ryder, and J. M. Yeomans. Transport coefficients of a mesoscopic fluid dynamics model. *J. Chem. Phys.*, 119:6388–6395, 2003.
34. N. Kikuchi, J. F. Ryder, C. M. Pooley, and J. M. Yeomans. Kinetics of the polymer collapse transition: The role of hydrodynamics. *Phys. Rev. E*, 71:061804, 2005.
35. A. Lamura and G. Gompper. Numerical study of the flow around a cylinder using multi-particle collision dynamics. *Eur. Phys. J. E*, 9:477–485, 2002.
36. A. Lamura, G. Gompper, T. Ihle, and D. M. Kroll. Multiparticle collision dynamics: Flow around a circular and a square cylinder. *Europhys. Lett.*, 56:319–325, 2001.
37. L. D. Landau and E. Lifshitz. *Statistical Physics*. Butterworth-Heinemann, 1980.
38. S. H. Lee and R. Kapral. Mesoscopic description of solvent effects on polymer synamics. *J. Chem. Phys.*, 124:214901, 2006.
39. M. P. Lettinga, J. K. G. Dhont, Z. Zhang, S. Messlinger, and G. Gompper. Hydrodynamic interactions in rod suspensions with orientational ordering. *Soft Matter*, 6:4556–4562, 2010.
40. D. Lüsebrink and M. Ripoll. Temperature inhomogeneities simulated with multiparticle-collision dynamics. *J. Chem. Phys.*, 136:084106, 2012.
41. D. Lüsebrink, M. Yang, and M. Ripoll. Thermophoresis of colloids by mesoscale simulations. *J. Phys.: Condens. Matter*, 24:284132, 2012.
42. A. Malevanets and R. Kapral. Mesoscopic model for solvent dynamics. *J. Chem. Phys.*, 110:8605–8613, 1999.
43. A. Malevanets and R. Kapral. Solute molecular dynamics in a mesoscopic solvent. *J. Chem. Phys.*, 112:7260–7269, 2000.
44. A. Malevanets and R. Kapral. Mesoscopic multi-particle collision model for fluid flow and molecular dynamics. *Lec. Notes Phys.*, 640:116, 2004.
45. A. Malevanets and J. M. Yeomans. Dynamics of short polymer chains in solution. *Europhys. Lett.*, 52:231–237, 2000.
46. G. R. McNamara and G. Zanetti. Use of the boltzmann equation to simulate lattice-gas automata. *Phys. Rev. Lett.*, 61:2332, 1988.
47. J. L. McWhirter, H. Noguchi, and G. Gompper. Flow-induced clustering and alignment of red blood cells in microchannels. *Proc. Natl. Acad. Sci.*, 106:6039–6043, 2009.
48. J. Morthomas and A. Würger. Hydrodynamic attraction of immobile particles due to interfacial forces. *Phys. Rev. E*, 81:051405, 2010.
49. F. Müller-Plathe. A simple nonequilibrium molecular dynamics method for calculating the thermal conductivity. *J. Chem. Phys.*, 106:6082–6085, 1997.
50. K. Mussawisade, M. Ripoll, R. G. Winkler, and G. Gompper. Polymer dynamics in a mesoscopic solvent. *J. Chem. Phys.*, 123:144905, 2005.
51. H. Noguchi and G. Gompper. Fluid vesicles with viscous membranes in shear flow. *Phys. Rev. Lett.*, 93:258102, 2004.
52. H. Noguchi and G. Gompper. Dynamics of fluid vesicles in shear flow: Effect of membrane viscosity and thermal fluctuations. *Phys. Rev. E*, 72:011901, 2005.
53. H. Noguchi, N. Kikuchi, and G. Gompper. Particle-based mesoscale hydrodynamic techniques. *Europhys. Lett.*, 78:10005, 2007.
54. J. T. Padding and A. A. Louis. Hydrodynamic and brownian fluctuations in sedimenting suspensions. *Phys. Rev. Lett.*, 93:220601, 2004.

55. J. T. Padding and A. A. Louis. Hydrodynamic interactions and brownian forces in colloidal suspensions: Coarse-graining over time and length-scales. *Phys. Rev. E*, 74:031402, 2006.
56. J. T. Padding, A. Wysocki, H. Löwen, and A. A. Louis. Stick boundary conditions and rotational velocity auto-correlation functions for colloidal particles in a coarse-grained representation of the solvent. *J. Phys.: Condens. Matter*, 17:S3393, 2005.
57. I. Pagonabarraga, M. H. J. Hagen, and D. Frenkel. Self-consistent dissipative particle dynamics algorithm. *Europhys. Lett.*, 42:377, 1998.
58. R. Piazza and A. Parola. Thermophoresis in colloidal suspensions. *J. Phys.: Condens. Matter*, 20:153102, 2008.
59. O. Pohl and H. Stark. Dynamic clustering and chemotactic collapse of self-phoretic active particles. *Phys. Rev. Lett.*, 112:238303, Jun 2014.
60. C. M. Pooley and J. M. Yeomans. Kinetic theory derivation of the transport coefficients of stochastic rotation dynamics. *J.Phys. Chem.B*, 109:6505–6513, 2005.
61. S. Y. Reigh, R. G. Winkler, and G. Gompper. Synchronization and bundling of anchored bacterial flagella. *Soft Matter*, 8:4363–4372, 2012.
62. M. Ripoll. Helicopter rotation and smectic-isotropic coexistence of strongly attractive rods. *Phys. Rev. E*, 83:040701(R), 2011.
63. M. Ripoll, P. Holmqvist, R. G. Winkler, G. Gompper, J. K. G. Dhont, and M. P. Lettinga. Attractive colloidal rods in shear flow. *Phys. Rev. Lett.*, 101:168302, 2008.
64. M. Ripoll, K. Mussawisade, R. G. Winkler, and G. Gompper. Low-reynolds-number hydrodynamics of complex fluids by multi-particle collision dynamics. *EPL*, 68:106–112, 2004.
65. M. Ripoll, K. Mussawisade, R. G. Winkler, and G. Gompper. Dynamic regimes of fluids simulated by multiparticle-collision dynamics. *Phys. Rev. E*, 72:016701, 2005.
66. M. Ripoll, R. G. Winkler, and G. Gompper. Star polymers in shear flow. *Phys. Rev. Lett.*, 96:188302, 2006.
67. M. Ripoll, R. G. Winkler, and G. Gompper. Hydrodynamic screening of star polymers in shear flow. *Eur. Phys. J. E*, 23:249–354, 2007.
68. G. Rückner and R. Kapral. Chemically powered nanodimers. *Phys. Rev. Lett.*, 98:150603, 2007.
69. S. P. Singh, R. G. Winkler, and G. Gompper. Nonequilibrium forces between dragged ultrasoft colloids. *Phys. Rev. Lett.*, 107:158301, 2011.
70. S. Succi. *The Lattice Boltzmann Equation: for fluid dynamics and beyond*. Clarendon, Oxford, 2001.
71. D. J. Tritton. *Physical Fluid Dynamics*. Oxford Science publications, second edition, 1988.
72. E. Tüzel, T. Ihle, and D. M. Kroll. Dynamic correlations in stochastic rotation dynamics. *Phys. Rev. E*, 74:056702, 2006.
73. E. Tüzel, G. Pan, T. Ihle, and D. M. Kroll. Mesoscopic model for the fluctuating hydrodynamics of binary and ternary mixtures. *Europhys. Lett.*, 80:40010, 2007.
74. E. Tüzel, M. Strauss, T. Ihle, and D. M. Kroll. Transport coefficients for stochastic rotation dynamics in three dimensions. *Phys. Rev. E*, 68:036701, 2003.
75. F. M. Weinert and D. Braun. Observation of slip flow in thermophoresis. *Phys. Rev. Lett.*, 101:168301, 2008.
76. S. Wiegand. Thermal diffusion in liquid mixtures and polymer solutions. *J. Phys.: Condens. Matter*, 16:R357–R379, 2004.

77. A. Würger. Thermophoresis in colloidal suspensions driven by marangoni forces. *Phys. Rev. Lett.*, 98:138301, 2007.
78. A. Würger. Thermal non-equilibrium transport in colloids. *Rep. Prog. Phys.*, 73:126601, 2010.
79. A. Wysocki, C. P. Royall, R. G. Winkler, G. Gompper, H. Tanaka, A. van Blaaderen, and H. Löwen. Direct observation of hydrodynamic instabilities in a driven non-uniform colloidal dispersion. *Soft Matter*, 5:1340–1344, 2009.
80. M. Yang, R. Liu, M. Ripoll, and K. Chen. A microscale thermophoretic turbine driven by external diffusive heat flux. *Nanoscale*, 6:13550–13554, 2014.
81. M. Yang and M. Ripoll. Simulations if thermophoretic nanoswimmers. *Phys. Rev. E*, 84:061401, 2011.
82. M. Yang and M. Ripoll. Drift velocity in non-isothermal inhomogeneous systems. *J. Chem. Phys.*, 136:204508, 2012.
83. M. Yang and M. Ripoll. Driving forces and polymer hydrodynamics in the soret effect. *J. Phys.: Condens. Matter*, 24:195101, 2012.
84. M. Yang and M. Ripoll. Thermophoretically induced flow field around a colloidal particle. *Soft Matter*, 9:4661, 2013.
85. M. Yang and M. Ripoll. A self-propelled thermophoretic microgear. *Soft Matter*, 10:1006, 2014.
86. M. Yang, M. Theers, J. Hu, G. Gompper, R. G. Winkler, and M. Ripoll. Effect of the angular momentum conservation on hydrodynamic simulations of colloids, 2015. (preprint).
87. M. Yang, A. Wysocki, and M. Ripoll. Hydrodynamic simulations of self-phoretic microswimmers. *Soft Matter*, 10:6208, 2014.
88. Y. Yang, V. Marceau, and G. Gompper. Swarm behavior of self-propelled rods and swimming flagella. *Phys. Rev. E*, 82:031904, 2010.
89. J. M. Yeomans. Mesoscale simulations: Lattice Boltzmann and particle algorithms. *Physica A*, 369:159–184, 2006.
90. A. Zöttl and H. Stark. Hydrodynamics determines collective motion and phase behavior of active colloids in quasi-two-dimensional confinement. *Phys. Rev. Lett.*, 112:118101, Mar 2014.

Responsive Particle Dynamics for Modeling Solvents on the Mesoscopic Scale

Wim Briels

Computational Biophysics, University of Twente
P.O. Box 217, 7500AE Enschede, The Netherlands
and
Forschungszentrum Jülich, ICS 3, D-52425 Jülich, Germany

E-mail: w.j.briels@utwente.nl

In this chapter we will review the standard ways to perform stochastic simulations on soft matter without memory. In order to make these methods available for very concentrated complex polymeric soft matter in which the particles experience large frictions we will extend these models to include memory and hydrodynamics at the Brownian level. This allows us to study flow instabilities and even flow in complex geometries.

1 Introduction

It is our intention in this paper to describe a model for simulating complex soft matter systems, possibly flowing through complex geometries, and possibly undergoing phase transitions or flow instabilities. This requires that the model faithfully represents all thermodynamic and all rheological properties of the system. The simplest way to do this would be to perform a full glory particle based simulation, if not this would be impossible because of prohibitively large computational costs (just to make an understatement). Yet this is more or less what we want to do, although we will have to remove the adjective 'full glory'.

Complex soft matter systems usually consist of very large particles with many internal degrees of freedom, dissolved in an appropriate solvent or in the molten state. The prototypical example of a particle is a star polymer of hundreds of kilo-Daltons. We will develop a severely coarse-grained simulation model for such a system, in which each molecule is represented by just its center of mass position. This requires special tricks of course, as can be understood from a glance at Fig. (1). In the left panel of this figure we have depicted an artist's impression of a star polymer solution. Each polymer consists of many arms connected to a central point. The concentration is such that the arms of each polymer are highly entangled with many of their neighbors. One of the characteristics of polymers as opposed to normal molecules is that they interact with hundreds of neighbors, not just something between ten and twenty as in dense liquids or solutions of normal molecules. Now, the act/art of coarse-graining is to remove all explicit reference to the internal degrees of freedom of each molecule and the degrees of freedom of the solvent. Once we have done this, our system looks geometrically like the right hand panel of Fig. (1). This looks very much like an ideal gas, yet we have to make the particles move like the centers of mass do in the the left hand panel.

We will gradually build the theory underlying the model that we need. We start by reviewing the theory behind the Langevin equation that describes the motion of colloidal

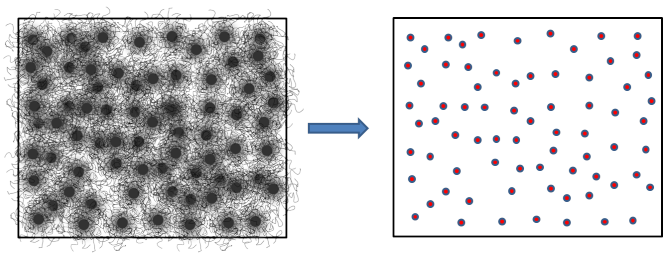


Figure 1: Definition of coarse-graining. In the left panel a complex soft matter system is depicted, consisting of star polymers dissolved in a solvent. Each star consists of many arms connected to a central point, while each arm is a polymer of many kilo Daltons of mass. The concentration is such that each star interacts with many, usually hundreds, of neighboring fellows. In the right panel we have depicted what the system looks like when only centers of mass are seen. This is what is simulated by our coarse model. The dynamics of the points in the right panel must be equal to that of the centers of mass in the left panel.

particles in a solvent. Next we discuss changes to be made in order to make the model applicable to our aimed for systems. In section 4 we notice that momenta usually play no role in soft matter systems and that it is better to dispense with them altogether. The result of this section will be the final model that we apply in our simulations. It describes displacements of the particles with respect to a background flow which drags them along. Our algorithm includes updates to calculate the time evolution of the background flow. Interactions between particles will be such that they give rise to correct thermodynamical behavior of the system. A reasonably small set of structural variables is introduced to describe memory effects that result from the fact that every configuration of the coarse degrees of freedom on the small scale may be accompanied by one of many configurations of the eliminated degrees of freedom. These variables describe how the eliminated dgrees of freedom would have responded to the changing coarse configurations. The method therefore carries the name of Resposive Particle Dynamics (RaPiD). It is important that the additional structural variables do not influence the thermodynamics of the system. Finally in section 5 we describe one example that has been simulated using this method.

2 The Langevin Equation

Our discussion in this section will be very much along the lines of McQuarrie's textbook on Statistical Physics[1]. The prototypical example of a system to which stochastic differential equations are applied is the infinitely diluted suspension of colloids in an appropriate solvent. In this case each colloidal particle can be assumed to move independently of all other particles, while only being influenced by the surrounding solvent molecules. In the usual explanation of the action of the solvent molecules on the colloid it is assumed, that while the colloid moves through the liquid, it experiences more collisions with the solvent molecules in the front than in its back. This gives rise, on average, to a force opposite to the direction of its motion and proportional to the absolute value of its velocity. The deviation of the instantaneous force from this average force is then a quickly fluctuating contribution, which on average equals zero. The corresponding equation of motion for the

colloid is called the Langevin equation and reads

$$m\frac{d\mathbf{v}}{dt} = -\xi\mathbf{v} + \mathbf{F}^{ran}. \tag{1}$$

Here $\mathbf{v}$ denotes the velocity of the particle and m its mass; ξ is called the friction coefficient and $\mathbf{F}^{ran}$ the random force. Notice that the dimensions of ξ are kg/s. Below we will prove that the random forces are related to the friction coefficient according to

$$< \mathbf{F}^{ran}(t)\mathbf{F}^{ran}(0) >= 2\xi k_B T\delta(t)\mathbf{1}. \tag{2}$$

The latter equation is called the Fluctuation-Dissipation theorem.

The left hand side of Eq. (2) should be read like

$$\begin{pmatrix} < F_x^{ran}(t)F_x^{ran}(0) > & < F_x^{ran}(t)F_y^{ran}(0) > & < F_x^{ran}(t)F_z^{ran}(0) > \\ < F_y^{ran}(t)F_x^{ran}(0) > & < F_y^{ran}(t)F_y^{ran}(0) > & < F_y^{ran}(t)F_z^{ran}(0) > \\ < F_z^{ran}(t)F_x^{ran}(0) > & < F_z^{ran}(t)F_y^{ran}(0) > & < F_z^{ran}(t)F_z^{ran}(0) > \end{pmatrix} \tag{3}$$

Pointy brackets in general indicate an equilibrium average. When we average the product of two variables taken at different times, the interpretation is a bit more involved. In the present case, $< F_\alpha^{ran}(t)F_\beta^{ran}(0) >$ is called the time correlation function of the random forces in α and β direction. It is the average value of $F_\alpha^{ran}(t)$ at time t over all realizations which had $F_\beta^{ran}(0)$ at time $t = 0$, multiplied by $F_\beta^{ran}(0)$ and next averaged over all possible realizations. A realization is just a time sequence of positions of the colloidal particle under investigation. A practical way to calculate $< F_\alpha^{ran}(t)F_\beta^{ran}(0) >$ from a realization of length T is

$$< F_\alpha^{ran}(t)F_\beta^{ran}(0) >= \frac{1}{T-t}\int_0^{T-t} d\tau F_\alpha^{ran}(\tau+t)F_\beta^{ran}(\tau) \tag{4}$$

$\mathbf{1}$ is the 3×3 unit matrix. Moreover, $\delta(t)$ is the Dirac delta, which may be thought of as an infinitely narrow function centered at $t = 0$ and whose integral is equal to unity. For its numerical treatment see the section on Brownian Dynamics.

Next consider a dense system of N colloids which interact via a potential $\Phi(r^{3N})$, where $r^{3N} = \{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N\}$. A conservative force

$$\mathbf{F}_i = -\nabla_i\Phi \tag{5}$$

then acts on particle i. We assume that the Langevin equation for particle i may be obtained by adding a subscript i to all quantities and by adding $-\nabla_i\Phi$ in the right hand side of Eq. (1). Besides this we assume that the frictions and random forces are not influenced by these extensions. The final equations then read:

$$m_i\frac{d\mathbf{v}_i}{dt} = -\nabla_i\Phi - \xi_i\mathbf{v}_i + \mathbf{F}_i^{ran}, \tag{6}$$

$$< \mathbf{F}_i^{ran}(t)\mathbf{F}_j^{ran}(0) >= 2\xi_i k_B T\delta(t)\delta_{ij}\mathbf{1}. \tag{7}$$

Here we have assumed that random forces acting on different particles are uncorrelated, which is reflected in the Kronecker delta δ_{ij} in the right hand side of the Fluctuation-Dissipation theorem, which equals unity when $i = j$ and zero otherwise.

Let me finally notice that we have ignored the possibility that moving one particle has an influence on other particles via the flow field induced in the solvent. The corresponding

interactions are called hydrodynamic interactions. They are indeed relevant for colloidal suspensions in Newtonian fluids, but play a minor role in the applications of stochastic methods presented in his chapter.

2.1 Fluctuation Dissipation theorem

In this section we restrict ourselves again to particles moving independently from each other in absence of any external fields. It will be assumed that friction forces and corresponding random forces, resulting from interactions with the solvent, are independent of possible interactions between the particles. Only in the case of hydrodynamic interactions a more general treatment is needed.

In order to prove the Fluctuation-Dissipation theorem we notice that

$$\mathbf{v}(t) = \mathbf{v}(0)e^{-\xi t/m} + \frac{1}{m}\int_0^t dt' e^{-\xi(t-t')/m}\mathbf{F}^{ran}(t') \tag{8}$$

solves the Langevin equation for the initial velocity $\mathbf{v}(0)$. We next invoke the equipartition theorem of statistical physics, which says that, in equilibrium, velocities along different Cartesian axes are uncorrelated, *i.e.* that $< v_\alpha(t)v_\beta(t) >=< v_\alpha(t) >< v_\beta(t) >= 0$ in case $\alpha \neq \beta$, while $< v_\alpha(t)v_\alpha(t) >= k_BT/m$. In order to make the particle forget its initial velocity and to reach equilibrium, we must take the limit of $t \longrightarrow \infty$ in the above solution. The equipartition theorem then says

$$\lim_{t\to\infty} < \mathbf{v}(t)\mathbf{v}(t) >= \frac{k_BT}{m}\mathbf{1}. \tag{9}$$

Inserting Eq. (8) we obtain

$$\lim_{t\to\infty} \frac{1}{m^2}\int_0^t dt' \int_0^t dt'' e^{-\xi(2t-t'-t'')/m} < \mathbf{F}^{ran}(t')\mathbf{F}^{ran}(t'') >= \frac{k_BT}{m}\mathbf{1}. \tag{10}$$

Next we make use of $< \mathbf{F}^{ran}(t')\mathbf{F}^{ran}(t'') >=< \mathbf{F}^{ran}(t''-t')\mathbf{F}^{ran}(0) >$ and change integration variables to $\tau' = t'' + t'$ and $\tau'' = t'' - t$, obtaining

$$\lim_{t\to\infty} \frac{1}{2m^2}\int_0^{2t} d\tau' \int_{-t}^{t} d\tau'' e^{-\xi(2t-\tau')/m} < \mathbf{F}^{ran}(\tau'')\mathbf{F}^{ran}(0) >= \frac{k_BT}{m}\mathbf{1}. \tag{11}$$

Performing the integral over τ' we obtain

$$\int_{-t}^{t} d\tau'' < \mathbf{F}^{ran}(\tau'')\mathbf{F}^{ran}(0) >= 2\xi k_BT\mathbf{1}. \tag{12}$$

If we next assume that the time-correlation functions of the random forces are very short lived, expressed as $< F_i^{ran}(t)F_i^{ran}(0) >= C\delta(t)$, and perform the final integral, we obtain the Fluctuation-Dissipation theorem.

2.2 Einstein equation

We now integrate the equations of motion once more in order to obtain displacements. Also in this subsection we ignore effects of interactions and external fields on the displacements of the particles and concentrate on the contributions from the solvent.

Integrating Eq. (8) we obtain

$$\mathbf{r}(t) - \mathbf{r}(0) = \mathbf{v}(0)\frac{m}{\xi}(1 - e^{-\xi t/m}) + \frac{1}{m}\int_0^t dt' \int_0^{t'} dt'' e^{-\xi(t'-t'')/m}\mathbf{F}^{ran}(t''). \quad (13)$$

By interchanging integrations we simplify the second term in the right hand side, obtaining

$$\frac{1}{m}\int_0^t dt'' \int_{t''}^t dt' e^{-\xi(t'-t'')/m}\mathbf{F}^{ran}(t'') = \frac{1}{\xi}\int_0^t dt''(1 - e^{-\xi(t-t'')/m})\mathbf{F}^{ran}(t''). \quad (14)$$

Using the Fluctuation-Dissipation theorem it is now a simple task to calculate

$$\begin{aligned} < \Delta\mathbf{r}(t)\Delta\mathbf{r}(t) > = & \frac{m^2}{\xi^2}(1 - e^{-\xi t/m})^2\mathbf{v}(0)\mathbf{v}(0) \\ & + \frac{k_B T}{\xi}(2t - 3\frac{m}{\xi} + 4\frac{m}{\xi}e^{-\xi t/m} - \frac{m}{\xi}e^{-2\xi t/m})\mathbf{1}, \end{aligned} \quad (15)$$

where $\Delta\mathbf{r}(t) = \mathbf{r}(t) - \mathbf{r}(0)$. Usually we are interested in equilibrium situations, where we may average the first term over the Maxwell distribution, obtaining $< \mathbf{v}(0)\mathbf{v}(0) >= \frac{k_B T}{m}\mathbf{1}$ and

$$< \Delta\mathbf{r}(t)\Delta\mathbf{r}(t) >= 2\frac{k_B T}{\xi}(t - \frac{m}{\xi} + \frac{m}{\xi}e^{-\xi t/m})\mathbf{1}. \quad (16)$$

For times t much larger than $\frac{m}{\xi}$ we may neglect the last two terms. We then calculate the mean square displacement by taking the trace

$$< \Delta\mathbf{r}(t) \cdot \Delta\mathbf{r}(t) >= Tr < \Delta\mathbf{r}(t)\Delta\mathbf{r}(t) >= 6Dt, \quad (17)$$

where

$$D = \frac{k_B T}{\xi}. \quad (18)$$

D is called diffusion coefficient and Eq. (18) the Einstein relation.

3 Application to Complex Soft Matter

In this section we will address the various terms in the Langevin equation from the perspective of its application to complex, visco-elastic soft matter systems. The prototypical system to have in mind is that of a melt or concentrated solution of highly branched polymers. Our intention is to reduce the degrees of freedom of each polymer to include only its three positional degrees of freedom. This means that all solvent molecules and all internal degrees of freedom of the polymers are eliminated from our description and can only influence the dynamics of the polymers via the three terms in the right hand side of the Langevin equation. For more information, see Briels[2] and van den Noort *et. al.*[3].

3.1 Potential of mean force

The main difference between simple molecular fluids and polymer systems is that simple molecules are rather compact objects packed in a liquid such that they interact with about ten to twenty neighboring fellow molecules, while polymers are large open structures interacting with hundreds of other polymers. As a result the usual assumption of pair wise additive interactions will not be applicable to the case of highly coarse-grained polymer solutions or melts.

In order to get a better understanding of the interactions in polymer systems, notice that $-\nabla_i \Phi$ is the average force on coarse particle i in case all velocities are equal to zero, *i.e.* for a fixed configuration $r^{3N} = \{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N\}$. Denoting the potential energy of the fully detailed microscopic system as $V(r^{3N}, q^M)$, where q^M denotes all the eliminated degrees of freedom, the average force on particle i reads

$$\mathbf{F}_i = -\frac{\int dq^M \nabla_i V(r^{3N}, q^M) \exp\{-\beta V(r^{3N}, q^M)\}}{\int dq^M \exp\{-\beta V(r^{3N}, q^M)\}}. \tag{19}$$

Here we have assumed that dq^M includes possible Jacobians necessary to provide the correct volume element. Comparison of this expression and $\mathbf{F}_i = -\nabla_i \Phi$ leads to

$$\Phi(r^{3N}) = -k_B T \ln \int dq^M \exp\{-\beta V(r^{3N}, q^M)\}. \tag{20}$$

This says that $\Phi(r^{3N})$ is the free energy of the eliminated degrees of freedom, *i.e.* the q^M, in the presence of the field provided by the retained coordinates r^{3N}. Now with Φ being a free energy we write

$$\Phi(r^{3N}) = \sum_{i=1}^{N} a(\eta_i(r^{3N})), \tag{21}$$

where $a(\eta_i(r^{3N}))$ is the free energy per polymer in a system with polymer volume fraction equal to the local volume fraction $\eta_i(r^{3N})$, with the contribution of the center of mass excluded. The local volume fraction may be defined as

$$\eta_i(r^{3N}) = \frac{1}{\rho_{max}} \sum_{j=1}^{N} w(r_{ij}), \tag{22}$$

where ρ_{max} is the maximum polymer (number) density and $w(r)$ some normalized, monotonously decaying weight function with a range on the order of two or three times the radius of gyration of the polymers; the normalization is such that $\int d^3 r w(r) = 1$.

A popular expression for the free energy per polymer (excluding the contribution of its center of mass) derives from the Flory-Huggins free energy[4].

3.2 Memory

One of the characteristics of complex polymer solutions or melts is their visco-elasticity. In order to include the effects of visco-elasticity, the Langevin equation may be generalized to read

$$m_i \frac{d\mathbf{v}_i}{dt} = -\nabla_i \Phi - \int_0^t dt' \mathbf{v}_i(t') \cdot \boldsymbol{\zeta}_i(t - t'; r^{3N}(t')) + \mathbf{F}_i^{ran}, \tag{23}$$

$$< \mathbf{F}_i^{ran}(t)\mathbf{F}_j^{ran}(0) >= k_B T \boldsymbol{\zeta}_i(t; r^{3N}(0))\delta_{ij}. \tag{24}$$

Here $\boldsymbol{\zeta}_i$ is a 3×3 matrix which is multiplied into the vector $\mathbf{v}_i$. Besides this, we have assumed that the total friction force on particle i depends on its velocity at all times t' prior to t and that the corresponding friction tensor $\boldsymbol{\zeta}_i$ depends on the configuration r^{3N} time t'. The generalized friction is said to include 'memory'. We will not prove the Fluctuation-Dissipation theorem as given in Eq. (24). Notice however, that if we assume $\boldsymbol{\zeta}_i(t; r^{3N}(0)) = 2\xi_i\delta(t)\mathbf{1}$ we get back Eqs (6) and (7). Note also that, again, we neglect hydrodynamic interactions.

Let us now try to understand Eq. (23). Within the present context it is useful to rewrite the memory term to read

$$-\int_0^t d\mathbf{r}_i(t') \cdot \boldsymbol{\zeta}_i(t - t'; r^{3N}(t')). \tag{25}$$

This equation reveals the possibility of an entirely different interpretation of the friction term than the one used at the beginning of this chapter to motivate the original Langevin equation Eq. (1). It says that if we displace a particle by $d\mathbf{r}_i(t')$, at all times t after t' the particle will experience a force opposing the displacement and being linear in its components. The strength of the force will gradually fade away with increasing time lapse $t - t'$ after the displacement. The total force at time t is simply the sum of all contributions from displacements in the past.

From a computational point of view Eqs (23) and (24) constitute a very complicated set of equations. Not only do we need to keep track of a usually long history of configurations in order to calculate the friction tensor, but also do we have to sample random forces from complicated coupled distributions to guarantee that the Fluctuation-Dissipation condition Eq. (24) is met. We would rather prefer to deal with the so called Markovian situation where the friction depends only on the instantaneous configuration. Let us therefore investigate a bit closer the meaning of the friction term.

To simplify our considerations we concentrate on one particular displacement $d\mathbf{r}_i(t')$. As a result of this displacement the environment of the particle, in a more detailed treatment described by the eliminated degrees of freedom q^M, will be slightly perturbed to a non-equilibrium configuration. This leads to a force that is slightly different from the average force $-\nabla_i\Phi$ that goes with the new configuration. The difference is what is described by $-d\mathbf{r}_i(t') \cdot \boldsymbol{\zeta}_i(t - t'; r^{3N}(t'))$. With increasing time t it gradually fades away because the eliminated degrees of freedom relax to their new equilibrium state. So, friction in a visco-elastic material basically results from non-equilibrium of the eliminated coordinates.

The above strongly suggests that we restore the slowly evolving part of the friction force, *i.e.* its memory part, back into the conservative force that derives from the free energy Φ. The fast evolving part of the friction force may then be included as a Markovian friction force. To this end we introduce structural parameters $\lambda^m = \{\lambda_1, \lambda_2, ..., \lambda_m\}$, which together describe the thermodynamic state of the eliminated degrees of freedom. At equilibrium they take values $\lambda_{eq}^m(r^{3N}) = \{\lambda_1^{eq}(r^{3N}), \lambda_2^{eq}(r^{3N}), ..., \lambda_m^{eq}(r^{3N})\}$. We next assume that the conservative potential Φ may be replaced by $A(r^{3N}, \lambda^m) = \Phi(r^{3N}) + \Phi^t(r^{3N}, \lambda^m)$ with

$$\Phi^t(r^{3N}, \lambda^m) = \frac{1}{2}\sum_{s=1}^{m} \alpha_s(\lambda_s - \lambda_s^{eq}(r^{3N}))^2. \tag{26}$$

We assume that the λ_s have no dimensions, which implies that the α_s have dimensions of energy. Φ^t is called the transient potential, and the corresponding forces $-\nabla_i \Phi^t$ are called transient forces. For brevity of notation we define

$$A(r^{3N}, \lambda^m) = \Phi(r^{3N}) + \Phi(r^{3N}, \lambda^m) \tag{27}$$

and call it (total) free energy.

3.3 Flow

An obvious short-coming of the Langevin equation presented so far is the fact that the possibility of a moving background is not accounted for. Obviously the friction felt by a particle results not from its absolute motion, but from its velocity with respect to the velocity of the background, *i.e.* the average local velocity of the eliminated degrees of freedom. This can easily be incorporated into the Langevin equation by replacing $-\xi_i \mathbf{v}_i$ by $-\xi_i[\mathbf{v}_i - \mathbf{V}(\mathbf{r}_i)]$, so

$$m_i \frac{d\mathbf{v}_i}{dt} = -\nabla_i A - \xi_i[\mathbf{v}_i - \mathbf{V}(\mathbf{r}_i)] + \mathbf{F}_i^{ran}. \tag{28}$$

This addition makes it necessary to have a model to estimate the local background velocity $\mathbf{V}(\mathbf{r}_i)$ at the position of the particle $\mathbf{r}_i$. We will present a possible way to do this in the section on Brownian Dynamics.

It is appropriate at this point to mention a different way to deal with flowing systems. This is done in Dissipative Particle Dynamics (DPD) by introducing pairwise frictions, which automatically conserve momenta and give rise to a Galilei invariant algorithm[5]. We will not discuss this method here, as it cannot be easily adapted to the case of overdamped systems, *i.e.* systems with very large friction forces.

4 Brownian Dynamics

In many applications of stochastic simulations to soft matter the friction coefficient ξ_i is very large, which we call the overdamped case. In these cases the time step becomes limited, not by the rapidity with which the conservative forces change, but by the contributions of the random forces which become very large.

This can be seen from a naive integration of the Langevin equation:

$$m_i d\mathbf{v}_i = -\nabla_i A dt - \mathbf{v}_i \xi_i dt + \sqrt{\frac{2k_B T \xi_i}{dt}} \mathbf{\Theta}_i dt. \tag{29}$$

Here $\mathbf{\Theta}_i$ is a random vector characterised by $< \Theta_{i,\alpha} >= 0$ and

$$< \Theta_{i,\alpha} \Theta_{i,\beta} >= \delta_{\alpha\beta}. \tag{30}$$

Notice that with the last term in Eq. (29) we have represented the random force $\mathbf{F}_i^{ran}$ by $\sqrt{2k_B T \xi_i / dt}\mathbf{\Theta}_i$. Obviously this equals zero on average, while

$$< F_{i,\alpha}^{ran} F_{i,\beta}^{ran} >= \frac{2k_B T \xi_i}{dt} \delta_{\alpha\beta}. \tag{31}$$

Indeed, when dealing with discretized time, $\delta(t) = 1/dt$, so the last equation agrees with the Fluctuation-Dissipation theorem. Now, looking at Eq. (29) we notice that ξ_i and dt

always appear in the combination $\xi_i dt$. In particular in the random term this may be very annoying because it forces us to use small time steps in order to prevent large velocity changes, which lead to large displacements and finally to wildly fluctuating conservative forces. These problems may be circumvented by eliminating velocities altogether from the description. This we will do in the next subsection. Of course, if we are interested in flow, we will have to come up with a method to measure flow velocities. This we will do in the third subsection of this section.

4.1 Brownian propagator

This section is essentially a simplified version of a paper by Ermak and McCammon[6]. Suppose we are treating a system with very large friction. Characteristic for such systems is that, time intervals Δt exist such that, independent of the velocity at the beginning of the interval, the particle mostly samples velocities from a Maxwell distribution, while at the same time it hardly displaces. By the latter we mean that displacements during time Δt are such that the potential hardly changes. We may therefore choose time intervals Δt such that

$$\frac{1}{\Delta t}\int_t^{t+\Delta t} dt' m_i \frac{d\mathbf{v}_i}{dt'} = \frac{1}{\Delta t}[m_i \mathbf{v}_i(t+\Delta t) - m_i \mathbf{v}_1(t)] \tag{32}$$

is arbitrarily small, while at the same time

$$\frac{1}{\Delta t}\int_t^{t+\Delta t} dt' \nabla_i A(r^{3N}(t')) \approx \nabla_i A(r^{3N}(t)). \tag{33}$$

Let us have a closer look at the meaning of the approximation in the last equation.

Taylor-expanding $\nabla A(r^{3N}(t'))$ we find that the first term that is neglected in Eq. (33) is

$$\sum_j \nabla_i \nabla_j A(r^{3N}(t)) \cdot \frac{1}{\Delta t}\int_t^{t+\Delta t} dt' (\mathbf{r}_j(t') - \mathbf{r}_j(t)). \tag{34}$$

Since we are considering the overdamped case, the dominant contribution to $|\mathbf{r}_j(t') - \mathbf{r}_j(t)|$ is proportional to $(t'-t)^{1/2}$, all other contributions being proportional to $(t'-t)^n$ with $n \geq 1$. After integration over t' and division by Δt this is proportional to $\sqrt{\Delta t}$. All other contributions are proportional to larger powers of Δt. Our approximation therefore implies that we neglect all terms proportional to any power of Δt larger than zero.

Summarizing our results so far, we have

$$0 = -\nabla_i A - \frac{1}{\Delta t}\int_t^{t+\Delta t} dt' \xi_i(t') \mathbf{v}_i(t') + \frac{1}{\Delta t}\int_t^{t+\Delta t} dt' \mathbf{F}_i^{ran}(t') \tag{35}$$

Treating the last two terms similarly to the gradient term, we must evaluate them up to zeroth order in Δt. We start with

$$\frac{1}{\Delta t}\int_t^{t+\Delta t} dt' \xi_i(t') \mathbf{v}_i(t') = \frac{1}{\Delta t}\int_t^{t+\Delta t} dt' [\xi_i(t) + \sum_{j=1}^{N} \nabla_j \xi_i(t) \cdot (\mathbf{r}_j(t') - \mathbf{r}_j(t))] \mathbf{v}_i(t') \tag{36}$$

Performing the integration in the first term and rewriting the second, we obtain

$$\frac{\Delta \mathbf{r}_i(t)}{\Delta t}\xi_i(t) + \sum_{j=1}^{N} \nabla_j \xi_i(t) \cdot \frac{1}{\Delta t}\int_t^{t+\Delta t} dt'(\mathbf{r}_j(t') - \mathbf{r}_j(t))\mathbf{v}_i(t'), \tag{37}$$

where $\Delta \mathbf{r}_i(t) = \mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t)$. Contributions from potential forces to the integral are at least proportional to $(\Delta t)^2$, so we can restrict attention to contributions from random terms. These are uncorrelated among different particles, so only terms with $i = j$ will contribute. We rewrite the argument of the remaining integral to get

$$\frac{\Delta \mathbf{r}_i(t)}{\Delta t}\xi_i(t) + \nabla_i \xi_i(t) \cdot \frac{1}{2\Delta t}\int_t^{t+\Delta t} dt' \frac{d}{dt'}[(\mathbf{r}_i(t') - \mathbf{r}_i(t))(\mathbf{r}_i(t') - \mathbf{r}_i(t))], \tag{38}$$

where only random displacements should be taken into account. Performing the integral and averaging over all possible realizations, we get

$$\nabla_i \xi_i(t) \cdot \frac{1}{2\Delta t} < (\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t))(\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t)) >= \frac{k_B T}{\xi_i}\nabla_i \xi_i(t), \tag{39}$$

where we have used Eq. (16).

Collecting terms so far we have

$$0 = -\nabla_i A - \frac{\Delta \mathbf{r}_i}{\Delta t}\xi_i - \frac{k_B T}{\xi_i}\nabla_i \xi_i + \frac{1}{\Delta t}\int_t^{t+\Delta t} dt' \mathbf{F}_i^{ran}(t') \tag{40}$$

Now, consider the last term. Assuming that the friction ξ_i may be taken constant during the time interval Δt, the integral of the random forces may be written as

$$\frac{1}{\Delta t}\int_t^{t+\Delta t} dt' \mathbf{F}_i^{ran}(t') = \frac{1}{\Delta t}\sqrt{2k_B T \xi_i dt}\sum_{k=1}^{\frac{\Delta t}{dt}} \mathbf{\Theta}_{i,k}. \tag{41}$$

First of all, notice that the three Cartesian directions may be treated independently from each other. Next, the sum of N random numbers with mean zero and variance equal to unity is a random number with mean zero and variance equal to N, so

$$\frac{1}{\Delta t}\sqrt{2k_B T \xi_i dt}\sum_{k=1}^{\frac{\Delta t}{dt}} \mathbf{\Theta}_{i,k} = \frac{1}{\Delta t}\sqrt{2k_B T \xi_i dt}\sqrt{\frac{\Delta t}{dt}}\mathbf{\Theta}_i. \tag{42}$$

Introducing this into Eq. (40) and rearranging the result we get after some simple algebra

$$\Delta \mathbf{r}_i = -\frac{1}{\xi_i}\nabla_i A \Delta t + k_B T \nabla_i \frac{1}{\xi_i}\Delta t + \sqrt{\frac{2k_B T \Delta t}{\xi_i}}\mathbf{\Theta}_i. \tag{43}$$

This concludes our derivation of the Brownian propagator.

From now on we do not need to discriminate Δt from dt anymore, so we will replace Δt by dt again. Clearly, the whole procedure that we have gone through for displacements may be repeated for changes in the structural parameters λ^m. This finally leads to the Brownian propagator

$$d\mathbf{r}_i = -\frac{1}{\xi_i}\nabla_i A dt + k_B T \nabla_i \frac{1}{\xi_i} dt + \sqrt{\frac{2k_B T dt}{\xi_i}}\mathbf{\Theta}_i \tag{44}$$

$$d\lambda_s = -\frac{1}{\alpha_s \tau_s}\frac{\partial A}{\partial \lambda_s} dt + \sqrt{\frac{2k_B T dt}{\alpha_s \tau_s}}\Theta_s \tag{45}$$

where we have written the 'friction coefficient' that goes with λ_s as $\alpha_s\tau_s$. So, τ_s is a characteristic time governing the relaxation of λ_s towards its equilibrium value $\lambda_s^{eq}(r^{3N})$. We have assumed that τ_s does not depend on λ_s.

Notice that, whereas time step and friction always occurred as a product ξdt in the Langevin equation, they appear as a fraction $\frac{dt}{\xi}$ in the Brownian propagator. This means that with increasing frictions increasingly larger time steps can be used.

4.2 Memory

It is instructive[7] to formally integrate Eq. (45) and put the result into the equation that describes the dynamics of positions, Eq. (44). When doing so, we will neglect the random contribution in Eq. (45). Using the expression for Φ^t in Eq. (26) and assuming that all λ_s are equal to zero at time zero, we obtain

$$-\nabla_i\Phi^t = \sum_{s=1}^{m}\alpha_s[\frac{1}{\tau_s}\int_0^t dt'\lambda_s^{eq}(r^{3N}(t'))e^{-(t-t')/\tau_s} - \lambda_s^{eq}(r^{3N}(t))]\nabla_i\lambda_s^{eq}(r^{3N}(t)). \tag{46}$$

There are many ways to rewrite this equation. We restrict ourselves to just one of them. We rewrite the integral in the right hand side according to

$$\frac{1}{\tau_s}\int_0^t dt'\int_0^{t'} dt''\frac{d\lambda_s^{eq}}{dt''}e^{-(t-t')/\tau_s} = \frac{1}{\tau_s}\int_0^t dt''\int_{t''}^t dt'\frac{d\lambda_s^{eq}}{dt''}e^{-(t-t')/\tau_s}. \tag{47}$$

Next we perform the integral ove t' and write $\lambda_s^{eq}(t)$ as an integral of its derivative from time zero to time t, obtaining

$$-\nabla_i\Phi^t = -\sum_{s=1}^{m}\alpha_s\int_0^t dt'\frac{\partial\lambda_s^{eq}}{dt'}e^{-(t-t')/\tau_s}\nabla_i\lambda_s^{eq}(r^{3N}(t)). \tag{48}$$

Performing some cosmetic actions, this may be written in the suggestive form

$$-\nabla_i\Phi^t = -\sum_{j=1}^{N}\int_0^t d\mathbf{r}_j(t')\cdot[\sum_{s=1}^{m}\alpha_s\nabla_j\lambda_s^{eq}(r^{3N}(t'))\nabla_i\lambda_s^{eq}(r^{3N}(t))e^{-(t-t')/\tau_s}], \tag{49}$$

which has a strong resemblance to the generalized friction introduced earlier.

Let us now have a closer look at what type of variables can be used as structure parameters λ_s. We will illustrate our discussion again using the example of star polymers. The left most part of Fig. (2) shows two stars equilibrated at a distance r_{ij}. Next, these two stars are quickly displaced to a new distance r_{ij}^{new}. It is clearly seen that the arms of the stars are 'torn apart' and will need some time to adjust to the new situation. As time passes by, the arms more and more relax, until in the rightmost picture they have found the equilibrium configuration that goes with the new distance r_{ij}^{new}. We therefore define structure parameters λ_{ij} with every pair of particles (i,j) at distances smaller than some cutoff value R_c, not necessarily the same as in section (4.4) below. The corresponding equilibrium values $\lambda_{ij}^{eq}(r_{ij})$ depend on the distance r_{ij} between the two particles.

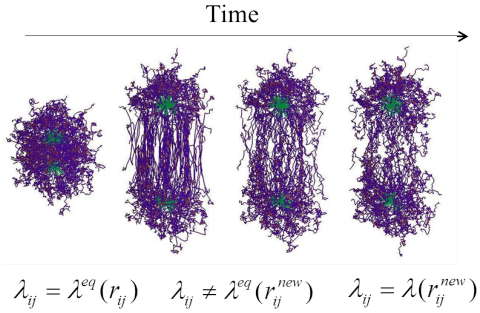


Figure 2: Two stars at equilibrium in the leftmost panel, displaced to a new distance in the second panel and next relaxing towards equilibrium at the new distance, which they reach in the rightmost panel. The thermodynamic state of the arms is described by a structural parameter λ_{ij} which equals $\lambda_{ij}^{eq}(r_{ij})$ in the leftmost panel and $\lambda_{ij}^{eq}(r_{ij}^{new})$ in the rightmost panel.

4.3 Equilibrium

We now ask what is the equilibrium distribution that goes with the Brownian propagator. We present our arguments for the case of a independent particles in a potential A. Generalization to the case of a dense suspension of particles is then obvious.

In general we may say that the probability $G(\mathbf{r})d^3r$ to find the particle in a cube of volume $d^3r = dxdydz$ evolves according to

$$\frac{\partial}{\partial t}G(\mathbf{r}) = -\nabla \cdot \mathbf{J}(\mathbf{r}), \tag{50}$$

where $\mathbf{J}$ is the flux of particles, *i.e.* the number of particles that go through unit area, per unit of time. Equilibrium occurs when the flux is equal to zero. The flux in our case may readily be understood to be given by

$$\mathbf{J}(\mathbf{r}) = -\frac{1}{\xi}G(\mathbf{r})\nabla A(\mathbf{r}) + k_B T G(\mathbf{r})\nabla\frac{1}{\xi(\mathbf{r})} + \mathbf{J}^{ran}(\mathbf{r}), \tag{51}$$

where $\mathbf{J}^{ran}(\mathbf{r})$ is the contribution from the random displacements made by the particles. Since random displacements in the three Cartesian directions are independent, we may restrict ourselves to treating just one of these.

Consider a plane of unit area at position x. The number of particles that pass through this plane during time Δt by making a step of size between Δ and $\Delta + d\Delta$ in positive x-direction is given by

$$\frac{1}{2}G(x-\Delta/2)P_{\Delta t}(x-\Delta/2;\Delta)\Delta d\Delta. \tag{52}$$

Here $P_{\Delta t}(x;\Delta)d\Delta$ is the probability that a particle at position x during time Δt makes a step of size between Δ and $\Delta + d\Delta$in positive x-direction. The factor of one half takes into account that only half of the particles move in positive direction. With a similar expression for particles passing through the plane in negative direction we obtain an overall contribution to J_x^{ran} from particles through the plane with step-size between Δ and $\Delta+d\Delta$

$$\frac{1}{2}G(x-\Delta/2)P_{\Delta t}(x-\Delta/2;\Delta)\Delta d\Delta - \frac{1}{2}G(x+\Delta/2)P_{\Delta t}(x+\Delta/2;\Delta)\Delta d\Delta. \tag{53}$$

After Taylor expanding G and $P_{\Delta t}$ to first order in Δ we obtain

$$-\frac{1}{2}\frac{\partial G(x)}{\partial x}P_{\Delta t}(x;\Delta)(\Delta)^2 d\Delta - \frac{1}{2}G(x)\frac{\partial}{\partial x}P_{\Delta t}(x;\Delta)(\Delta)^2 d\Delta. \tag{54}$$

We next integrate over all values of Δ and divide by Δt, obtaining

$$J_x^{ran} = -\frac{\partial G(x)}{\partial x}\frac{k_B T(x)}{\xi(x)} - G(x)\frac{\partial}{\partial x}\frac{k_B T(x)}{\xi}. \tag{55}$$

Generalizing to three Cartesian coordinates we obtain for the total flux

$$\mathbf{J}(\mathbf{r}) = -\frac{1}{\xi\mathbf{r})}G(\mathbf{r})\nabla A(\mathbf{r}) + k_B T(\mathbf{r})G(\mathbf{r})\nabla\frac{1}{\xi(\mathbf{r})} - \nabla[G(\mathbf{r})\frac{k_B T(\mathbf{r})}{\xi(\mathbf{r})}]. \tag{56}$$

We have included the possibility that the temperature depends on the position of the particle.

Now let us draw a few conclusions. First, rewrite the flux as

$$\mathbf{J}(\mathbf{r}) = -\frac{1}{\xi(\mathbf{r})}G(\mathbf{r})\nabla A(\mathbf{r}) - \frac{k_B T(\mathbf{r})}{\xi(\mathbf{r})}\nabla G(\mathbf{r}) - \frac{1}{\xi(\mathbf{r})}G(\mathbf{r})\nabla k_B T(\mathbf{r}). \tag{57}$$

We find that besides external forces, both concentration gradients and temperature gradients give rise to fluxes. Fluxes due to temperature gradients are called Soret fluxes.

Next, assume that the temperature is constant throughout the system. Putting the flux equal to zero, we find that the distribution is proportional to the Boltzmann factor, *i.e.*

$$G(\mathbf{r}) \propto e^{-A(\mathbf{r})/k_B T}. \tag{58}$$

So, the statistical equilibrium distribution gives rise to zero flux, as expected.

Finally, it is not difficult to understand that the equilibrium distribution that goes with Eqs (44,45) is given by

$$G(r^{3N}, \lambda^m) \propto e^{-A(r^{3N},\lambda^m)/k_B T}. \tag{59}$$

Integrating the last equation over all possible values of the λ_s we obtain

$$\begin{aligned} G(r^{3N}) &\propto \int d\lambda^m e^{-A(r^{3N},\lambda^m)/k_B T} \\ &= e^{-\Phi(r^{3N})/k_B T}\int d\lambda^m e^{-\frac{1}{2}\sum_{s=1}^{m}\alpha_s(\lambda_s-\lambda_s^{eq}(r^{3N}))^2/k_B T} \\ &\propto e^{-\Phi(r^{3N})/k_B T}. \end{aligned} \tag{60}$$

The distribution of configurations r^{3N} in the stationary state is therefore equal to the exact equilibrium distribution of the system (provided $\Phi(r^{3N})$ faithfully represents the free energy of the structural parameters as defined in Eq.(20)). As a result thermodynamic properties simulated with this model will be the exact thermodynamic properties of the system.

4.4 Flow

As they stand, our equations of motion, Eqs (44) and (45), are not Galilei invariant. In order to achieve this, we include an affine displacement $\mathbf{V}(\mathbf{r}_i)dt$ in the right hand side of Eq. (44). Notice that this term would have appeared automatically had we started our derivation of the Brownian propagator with a Langevin equation including a local velocity term. So its physical meaning is the same as in the Langevin equation. It is a background flow field with respect to which the particles move and experience friction forces.

Now this forces us to invent a way to calculate the local velocity $\mathbf{V}(\mathbf{r}_i)$ at the position of particle i. One way is to make use of a predefined velocity field obtained on the basis of some macroscopic phenomenological theory. If we want the flow to develop itself as a result of applied forces and boundary conditions, however, this will not do. We will assume that the background flow at position $\mathbf{r}$ may be obtained as some local average of the drift velocities of the particles near $\mathbf{r}$, drift velocities being defined as displacements defined by time intervals. Moreover, we assume that a good representation of the background flow field may be obtained by giving its value at a discrete set of points, for which we choose the positions of the particles. Defining $\mathbf{v}_i = \mathbf{V}(\mathbf{r}_i)$ it is possible to motivate the following update scheme for the local velocities $\mathbf{v}_i$

$$d\mathbf{v}_i = \frac{1}{\tau_f}[-\frac{1}{\xi_i}\nabla_i A + \sum_{j=1}^{N} f_{ij}(r_{ij})(\mathbf{v}_j - \mathbf{v}_i)]dt + \sum_{j=1}^{N}\sqrt{\frac{2k_BT}{\xi_i}f_{ij}}\frac{\mathbf{\Theta}_{ij}}{\tau_f}, \qquad (61)$$

where the random pair vectors must be such that $\mathbf{\Theta}_{ij} = -\mathbf{\Theta}_{ji}$. The functions f_{ij} are defined according to

$$f_{ij}(r_{ij}) = \frac{15}{2\pi R_c^3}\xi_j(\frac{1}{\rho_i} + \frac{1}{\rho_j})(1 - \frac{r_{ij}}{R_c})^3$$

$$\rho_i = \frac{21}{2\pi R_c^3}\sum_{j=1}^{N}\xi_j(1 - \frac{r_{ij}}{R_c})^4(4\frac{r_{ij}}{R_c} + 1). \qquad (62)$$

Here R_c is some cutoff distance such that on average about 15 particles contribute to the sum in ρ. We refer for further details to Padding and Briels[8].

When dealing with flowing matter, choosing the correct propagator is only part of the story. In case one is interested in the influence of hydrodynamics on equilibrium properties one can perform simulations with periodic boundary conditions. This has been done in the paper by Padding and Briels[8] and it was found that all theoretical results are reproduced correctly. If one is interested in flow instabilities that occur in the bulk of the system, non-equilibrium simulations with Lees-Edwards boundary conditions may be performed. This has been done to study shear banding with a variety of systems. The example of telechelic polymers will be discussed below. If, however, interest lies in systems with hard walls, one has to come up with the correct way to implement stick our partially stick boundary conditions. These methods have been developed for standard simulations, but for the present model are still under construction.

5 Telechelic Polymers

I will restrict myself to discussing one example of a simulation performed with the RaPiD scheme.

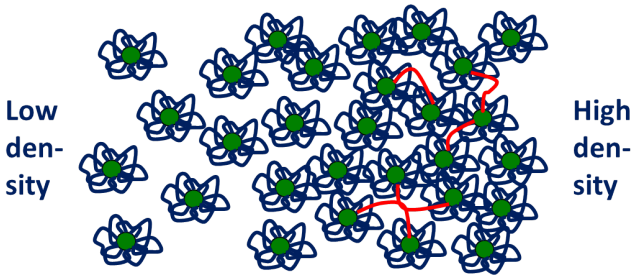


Figure 3: 3-block copolymers dissolved in water with increasing concentration from left to right. The two outer blocks are hyprophobic (green) and the middle block hydrophylic (blue). The hydrophobic middle blocks of some twenty polymers gather together to form micelles, while the middle blocks are dissolved in the solvent. At low concentrations micelles exist as independent flowers. At higher concentrations the two outer blocks of one polymer do not necessarily take part in one and the same micelle, but may be part of two different micelles, thereby forming bridges between them. Bridges are colored red for visualization purposes only.

In Fig. (3) a solution of 3-block co-polymers is depicted. In the left hand side of the figure the concentration is low, while in the right hand side it is large. The two outer blocks of the polymer are hydrophobic (green) while the inner block is hydrophylic (blue). At low concentrations the stars look like flowers. The outer, hydrophobic blocks of some twenty polymers have gathered together to form a micelle, thereby minimizing the unfavorable interaction with the solvent. The middle blocks are dissolved in the solvent. With increasing concentrations the hydrophonic outer blocks may take part in two different micelles, thereby establishing a bridge between the two micelles. These bridges severely influence the rheological properties of the fluid.

As mentioned above, the transient potential is assumed to be a pairwise sum of all pairs within a predescribed cutoff distance

$$\Phi^t(r^{3N}, \lambda^m) = \frac{1}{2} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \alpha(\lambda_{ij} - \lambda_{ij}^{eq}(r_{ij}))^2. \tag{63}$$

The meaning of the λ_{ij} ia taken to be the number of bridges between particles i and j. Then it is reasonable to assume that all α's are equal. Moreover the thermodynamic potential in this case is best represented by a sum of pair potentials as well. The pair contributions were calculated by means od Scheutjens-Fleer theory. All parameters in the model were known except α. For more information see Sprakel *et. al.*[9]. In Fig. (4) we present the results of our calculations of viscosities together with experimental values. The experimental viscosities were used to fit the only unknown parameter in the model. The interesting point is that viscosities vary by five orders of magnitude when the concentratios vary by three orders of magnitude, yet all of this can be reproduced by adjusting only one parameter.

Encouraged by our result so far, we now try to go for predictions. To this end we have performed non-equilibrium simulations in which the systems are sheared according to well established methods in simulation country. In Fig. (5) results are shown for a system with concentration of 20 gram/liter and a shear rate of $4/\tau$, where τ is the characteristic time occurring in the propagator for λ_{ij}. In the left panel, upper figure the stress is plotted as function of time lapse since the start of the run. It is clearly seen that, initially, the stress is constant, but, after some time, begins to drop until it reaches a new stationary value. In

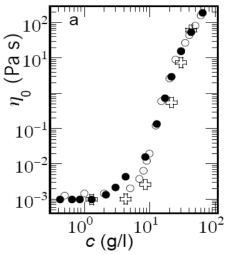


Figure 4: Viscosities versus concentration. Open circles are experimental values and crosses are theoretical results. The black circles are the results of simulations with artificially enlarged frictions. The reason for doing so is that these larger frictions allow larger time steps and therefore shorter simulations. As can be seen changing the friction to (incorrect) larger values only influences the results at small concentrations. This tells us that, while at low concentrations viscosities are determined by frictions with the solvent, at higher concentrations viscosities are determined by interactions between the particles, as expected.

the left panel, lower figure the flow velocity in x-direction is plotted for various equidistant positions along the gradient direction. In a normal situation, with a constant shear rate at all positions along the gradient direction, all lines would be parallel and equidistant (since the velocity would linearly increase along the gradient direction). We see however, that concomitantly with the drop of the stress velocities begin to change until after some time the velocities in a number of planes along the gradient direction differ from each other by only very small amounts. One may say that the system has split into two parts, one in which the shear rate is large, and one in which the shear rate is very low. In the example discussed here, the band with the lower shear rate has a somewhat larger concentration than the one with the higher shear rate. As a result the two bands can easily be visualized as seen in the right panel. The phenomenon just described is called shear banding, and has been theoretically analyzed by Dhont[10]. It is important to notice that banding would not have occurred in our simulations had we imposed a linear flow field instead of measured the background flow as described in the theory section.

In Fig. (6) we show some quantitative results for the banding observed with the system of Fig. (5). In the left panel we present the concentrations in the two bands as a function of applied shear rate. In cases when no banding occurs, the concentration is constant and equal to 20 gram/liter. For shear rates when banding does occur, two concentrations are given, one for each band. As can be seen they clearly differ. In the right panel, shear rates in the two bands are plotted as a function of the applied shear rate. Again, when no banding occurs, the system is homogeneously sheared with only one shear rate, equal to the applied shear rate. In this case, also experimental results are available (open circles), and it is seen that experiment and simulations are very well in agreement.

6 Concluding Remarks

In this paper we have reviewed the standard examples of stochastic dynamics simulations and adjusted them to be applicable to simulate the flow of complex soft matter in complex geometries. The changes needed to accommodate to soft matter systems consisted

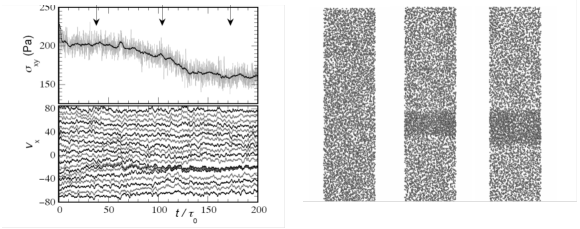


Figure 5: Results from non-equilibrium simulations with telechelic systems of concentrations equal to 20 gram/liter and shear rates equal to $4/\tau$. Left upper figure shows tresses as function of time lapse since the start of the run. After some initial period of constant stresses, the stresses begin to drop to lower values until they reach a new stationary value. Concomitantly with the stress drop the velocity field changes (left lower figure) such that the system splits into two bands, one of which has a shear rate smaller than the applied shear rate, while the other has a shear rate larger that the applied shear rate. Since the concentrations in the two bands differ, the bands can be easily observed, as shown in the right part of the figure. In these figures, flow is from left to right with increasing velocities from bottom to top. The ratio between flow velocity and vertical position is the shear rate. In the dark bands shear rates are less then in the lighter bands.

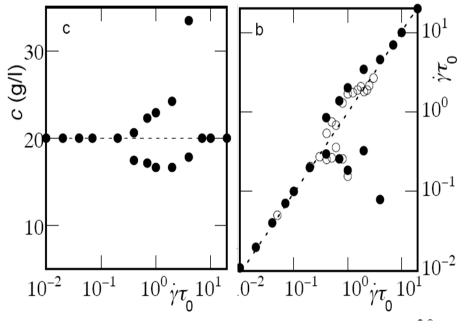


Figure 6: Quantitative results from the simulations described in Fig. (5). In the left panel the concentrations in the two bands are plotted as a function of the applied shear rates. When no banding occurs the concentration is constant throughout the box and equal to the overall concentration. When banding occurs the concentration in the high shear rate band is less than the overall concentration while that in the low shear rate band is larger than the overall concentration. In the right panel the corresponding shear rates are plotted as a function of the imposed shear rates. In this case experimental results are available and seen to be in very good agreement with the simulated results.

of inclusion of memory, adjusting the method to overdamped systems and guaranteeing hydrodynamics also for the Brownian propagator.

I have presented results for only one system, solutions of telechelic polymers. Several more systems have been successfully simulated by now. The linear rheology of linear polymers was shown to be well represented by the RaPiD model for frequencies just beyond the first crossing point of the shear and loss moduli. Non linear rheology was well reproduced up to rather high shear rates. One of the success stories of the method has been with pressure sensitive materials where both shear and elongational viscosities were well reproduced. In times where The Web of Science plays an important role in rating scientists

the interested reader may find results obtained with the RaPiD model by typing in any of the names mentioned in the Acknowledgements below.

Acknowledgments

I thank Johan Padding and Wouter den Otter for stimulating collaborations during more then a decade. Moreover I thank all students who worked on this model and contributed more than they will realize, among them Albert van den Noort, Peter Kindt, Igor Santos de Oliveira, Li Liu and Barry Fitzgerald.

References

1. D.A. McQuarrie, Statistical Mechanics, University Science Books, 2000 Sausalito, US.
2. W.J. Briels, Soft Matter **5**, 4401, 2009.
3. A. van den Noort, W.K. den Otter, and W.J. Briels, EPL **80**, 28003, 2007.
4. M. Doi, Introduction to Polymer Physics, Clarendon Press, 2009, Oxford, UK.
5. P. Espanol and P Warren, EPL **30**, 191, 1995.
6. D.L. Ermak and J.A. McCammon, J. Chem. Phys. **69**, 1352, 1978.
7. W.J. Briels, D. Vlassopoulos, K. Kang, and J.K.G. Dhont, J. Chem. Phys **134**, 124901, 2011.
8. J.T. Padding and W.J. Briels, J. Chem. Phys. **141**, 244108, 2014.
9. J. Sprakel, E. Spruijt, J. van der Gucht, J.T. Padding, and W.J. Briels, Soft Matter **5**, 4748, 2009.
10. J.K.G. Dhont, Phys. Rev. E **60**, 4534, 1999.

NFFT Based Fast Ewald Summation for Various Types of Periodic Boundary Conditions

Franziska Nestler, Michael Pippig and Daniel Potts

Technische Universität Chemnitz, Faculty of Mathematics
09107 Chemnitz, Germany

E-mail: {franziska.nestler, michael.pippig, daniel.potts}@mathematik.tu-chemnitz.de

The fast calculation of long-range interactions isno a demanding problem in particle simulation. In this tutorial we present fast Fourier-based methods for the Coulomb problem with mixed periodicity. The main focus of our approach is the decomposition of the problem into building blocks that can be efficiently realized. For that reason we recapitulate the fast Fourier transform at nonequispaced nodes (NFFT) and the fast summation method. Application of these two methods to the Ewald splitting formulas yields efficient methods for calculating the Coulomb energies in 3d-periodic, 2d-periodic, 1d-periodic, and also in 0d-periodic (open) boundary conditions.

1 Introduction

We start with a formal definition of the Coulomb problem with mixed periodic boundary conditions. Assume that N charges $q_j \in \mathbb{R}$ at positions $\boldsymbol{x}_j \in \mathbb{R}^3$, $j = 1, \ldots, N$, fulfill the charge neutrality condition

$$\sum_{j=1}^{N} q_j = 0. \tag{1}$$

The total Coulomb energy of the particle system can be formally written as

$$U_{\mathcal{S}} := \frac{1}{2} \sum_{j=1}^{N} q_j \phi_{\mathcal{S}}(\boldsymbol{x}_j),$$

where for each particle j the potential $\phi_{\mathcal{S}}(\boldsymbol{x}_j)$ is given by

$$\phi_{\mathcal{S}}(\boldsymbol{x}_j) := \sum_{\boldsymbol{n} \in \mathcal{S}} {\sum_{i=1}^{N}}' \frac{q_i}{\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|}. \tag{2}$$

Thereby, we denote by $\|\cdot\|$ the Euclidean norm and define the difference vectors $\boldsymbol{x}_{ij} := \boldsymbol{x}_i - \boldsymbol{x}_j$. The edge length of the simulation box in each dimension subject to periodic boundary conditions is given by $L > 0$. Furthermore, the set of translation vectors $\mathcal{S} \subseteq \mathbb{Z}^3$ will be defined later on according to the given boundary conditions. Note that the prime on the double sum indicates that for $\boldsymbol{n} = \boldsymbol{0}$ all terms with $i = j$ are omitted. We are also interested in the forces acting on the particles, which are given by

$$\boldsymbol{F}_{\mathcal{S}}(\boldsymbol{x}_j) := q_j \boldsymbol{E}_{\mathcal{S}}(\boldsymbol{x}_j), \quad \text{with the fields} \quad \boldsymbol{E}_{\mathcal{S}}(\boldsymbol{x}_j) := -\nabla \phi_{\mathcal{S}}(\boldsymbol{x}_j).$$

For the sake of brevity, we will derive fast algorithms for computing the potentials $\phi_{\mathcal{S}}(\boldsymbol{x}_j)$ and skip the analog derivation of algorithms for computing the forces $\boldsymbol{F}_{\mathcal{S}}(\boldsymbol{x}_j)$ within this tutorial.

The different cases of mixed periodic boundary conditions are described as follows. Assume periodic boundary conditions in the first $p \in \{0, 1, 2, 3\}$ dimensions and non-periodic (open) boundary conditions in the remaining $3 - p$ dimensions. Then, we set $\mathcal{S} := \mathbb{Z}^p \times \{0\}^{3-p}$ with $\boldsymbol{x}_j \in [-L/2, L/2)^p \times \mathbb{R}^{3-p}$, i.e., the sum over $\mathcal{S}$ in (2) can be interpreted as a replication of the primary box along all dimensions subject to periodic boundary conditions. For a graphical illustration see Figures 1 and 2.

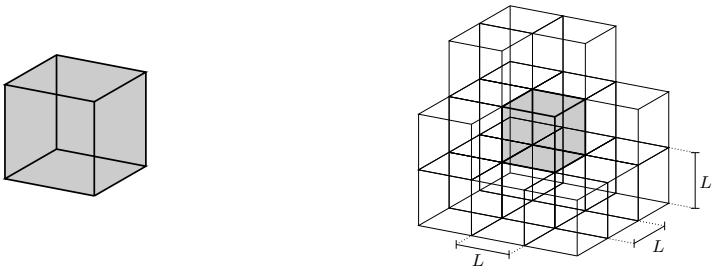


Figure 1: In the 0d-periodic case the particles are distributed within a finite box in $\mathbb{R}^3$ (left). In the 3d-periodic case the simulation box with edge length L is duplicated along all three dimensions (right).

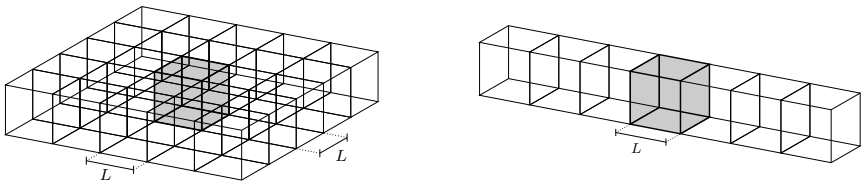


Figure 2: The simulation box is duplicated along two of three dimensions in the 2d-periodic case (left) and along one dimension in the 1d-periodic setting (right).

It is important to note that except for open boundary conditions the sum (2) is only conditionally convergent, i.e., the values of the potentials $\phi_{\mathcal{S}}(\boldsymbol{x}_j)$ depend on the order of summation. A common definition is to sum up the interactions box wise in a spherically increasing order, i.e.,

$$\phi_{\mathcal{S}}(\boldsymbol{x}_j) := \sum_{t=0}^{\infty} \sum_{\substack{\boldsymbol{n} \in \mathcal{S} \\ \|\boldsymbol{n}\|^2 = t}} \sum_{i=1}^{N}{}' \frac{q_i}{\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|}. \tag{3}$$

The well known Ewald summation technique[16] is the main basis for a variety of fast algorithms for the evaluation of (2) under 3d-periodic boundary conditions, see[26, 15, 11, 13, 33]. It is based on the trivial identity

$$\frac{1}{r} = \frac{\operatorname{erfc}(\alpha r)}{r} + \frac{\operatorname{erf}(\alpha r)}{r}. \tag{4}$$

Hereby, $\alpha > 0$ is generally known as the splitting parameter, $\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2} \mathrm{d}t$ is the error function and $\operatorname{erfc}(x) := 1 - \operatorname{erf}(x)$ is the complementary error function.

If (4) is applied to (2), the potential $\phi_{\mathcal{S}}(\boldsymbol{x}_j)$ is split into two parts. Thereby, the sum containing the erfc-terms includes a singularity at $r = 0$ but converges that fast that a good approximation is obtained by only considering few summands. The second part, containing the error function, is still conditionally convergent but exclusively involves smooth and L-periodic functions. The well known Ewald approach transforms this part into a fast convergent Fourier space sum under the implicit assumption of the spherical summation order (3). For a derivation in the 3d-periodic case we refer to the paper[12], where convergence factors are applied in order to calculate conditional convergent sum. A similar derivation of the Ewald formulas for the 2d- and 1d-periodic settings can be found in the Appendix of[40].

In the case of 3d-periodic boundary conditions the nonequispaced fast Fourier transform (NFFT)[30] can be directly applied to the Fourier space sum in order to achieve a fast algorithm. For all other kinds of mixed boundary conditions it is also possible to derive fast algorithms based on the NFFT. However, dimensions subject to non-periodic boundary conditions require special treatment in order to get fast convergent Fourier approximations. More precisely, we must embed non-periodic functions into smooth periodic functions, such that their Fourier sum converges rapidly, see Section 2.1 for details.

The outline of this tutorial is as follows. We start with some preliminary remarks about Fourier approximations and give a short introduction to the nonequispaced fast Fourier transform (NFFT) in Section 2. In Section 3 we present the main ideas of the fast Ewald summation for 3d-periodic boundary conditions. In Section 4 we consider the case of periodic boundary conditions in two of three dimensions. Thereby, we follow mainly the presentation from Section 4 in[40]. We continue in Section 5 with the 1d-periodic case in an analog manner as in Section 5 of[40]. Finally we extend the results to 0d-periodic (open) boundary conditions in Section 6. Finally, in Section 7 we conclude the tutorial and give references to numerical results.

2 Prerequisite

In this section we introduce three different concepts from Fouier analysis, which we apply in order to derive the presented algorithms, see Sections 3–6.

2.1 Fourier approximations

In the following, we discuss three different approaches to compute a Fourier approximation of a non-periodic function f within an interval $[-L, L]$.

Variant I (Periodization): The continuous Fourier transform of the function $f \in L^1(\mathbb{R})$ is given by

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x) \mathrm{e}^{-2\pi \mathrm{i} x \xi} \mathrm{d}x.$$

If f is sufficiently small outside the interval $[-L, L]$, we may approximate f by its h-periodic version $\sum_{n \in \mathbb{Z}} f(\cdot + hn)$, where $h \geq 2L$, apply the Poisson summation formula

and truncate the resulting infinite sum in order to obtain an approximation of the form

$$f(x) \approx \sum_{n=-\infty}^{\infty} f(x+hn) = \frac{1}{h} \sum_{l=-\infty}^{\infty} \hat{f}(l/h) \mathrm{e}^{2\pi \mathrm{i} l r/h}$$
$$\approx \frac{1}{h} \sum_{l=-M/2}^{M/2-1} \hat{f}(l/h) \mathrm{e}^{2\pi \mathrm{i} l r/h}, \tag{5}$$

where $M \in 2\mathbb{N}$ has to be chosen sufficiently large. Alternatively, we could argue as follows. First, we truncate the Fourier integral and, second, we approximate the resulting finite integral via the trapezoidal quadrature rule

$$f(x) = \int_{\mathbb{R}} \hat{f}(\xi) \mathrm{e}^{2\pi \mathrm{i} x \xi} \mathrm{d}\xi \approx \int_{-K/2}^{K/2} \hat{f}(\xi) \mathrm{e}^{2\pi \mathrm{i} r \xi} \mathrm{d}\xi$$
$$\approx \frac{K}{M} \sum_{l=-M/2}^{M/2-1} \hat{f}(\tfrac{lK}{M}) \mathrm{e}^{2\pi \mathrm{i} r l K/M}. \tag{6}$$

Comparison of (5) and (6) shows that this approach is equivalent to considering a $h = M/K$ periodization of f, as described above.

This approach is limited to functions that decay sufficiently fast in the interval $[-h/2, h/2)$. In other words, whenever f is not sufficiently small we need to choose a relatively large period $h \gg 2L$, which may also result in the choice of a large cutoff M.

Variant II (Truncation): We take a sufficiently large cutoff $h \geq 2L$ and approximate the function f on the interval $[-h/2, h/2]$ by a Fourier series

$$f(x) \approx \sum_{l=M/2}^{M/2-1} c_l \mathrm{e}^{2\pi \mathrm{i} l x/h},$$

where we compute the coefficients c_l by

$$c_l := \frac{1}{h} \int_{-h/2}^{h/2} f(x) \mathrm{e}^{-2\pi \mathrm{i} l x/h} \mathrm{d}x.$$

Note that the approximated h-periodic function is only smooth of order zero in $r = h/2$, which results in a rather slow second order convergence in Fourier space. Thus, one may have to choose M very large in order to achieve a good approximation. In contrast to Variant I, this approximation approach can be used for non decaying functions f as well.

Variant III (Regularization): Another approach to obtain a Fourier space representation of f is as follows. The key idea is to cutoff f outside the interval $[-L, L]$ but use a Fourier approximation on the slightly larger interval $[-h/2, h/2]$. In the resulting gap $[L, h-L]$ we construct a regularization function that interpolates the derivatives of f at L up to order $p-1 \in \mathbb{N}$. Therefore, we get a Fourier approximation of a $(p-1)$-times differentiable function which means $(p+1)$-th order convergence in Fourier space. In order to construct the smooth ($(p-1)$-times differentiable) transitions we have to regularize the function f. Thereby, we assume that we know the function values and the derivatives in the boundary points (in the following denoted by a_j and b_j) and compute a regularization (in the following denoted by P). The following theorem gives the precise definition of

the regularizing function. We remark that in our application we always know the function values and the derivatives in the boundary points. Methods without this knowledge are known as Fourier extensions[28] or Fourier continuations[35].

Theorem 2.1. *Let an interval $[m-r, m+r], r > 0$, and the interpolation values $a_j = f^{(j)}(m-r)$, $b_j = f^{(j)}(m+r)$, $j = 0, \dots, p-1$, be given. For $y = \frac{x-m}{r}$ the polynomial*

$$P(x) = \sum_{j=0}^{p-1} B(p,j,y) r^j a_j + \sum_{j=0}^{p-1} B(p,j,-y)(-r)^j b_j,$$

of degree $2p-1$, which is defined using the basis polynomials

$$B(p,j,y) := \sum_{k=0}^{p-1-j} \binom{p-1+k}{k} \frac{1}{j! 2^p 2^k} (1-y)^p (1+y)^{k+j},$$

satisfies the interpolation conditions $P^{(j)}(m-r) = a_j, P^{(j)}(m+r) = b_j, j = 0, \dots, p-1$.

Proof. See Corollary 2.2.6 in[3] or Proposition 3.2 in[17]. □

In summary, we see some graphical illustrations of the above-mentioned three Fourier approximation variants in Figures 3 and 4.

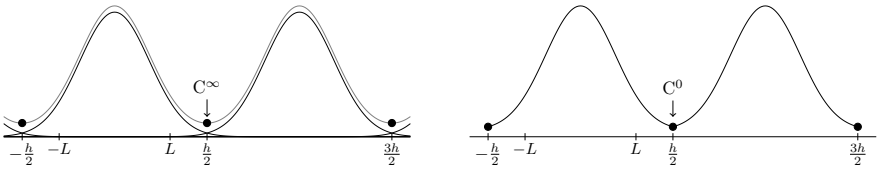


Figure 3: Variant I (periodization) on the left and Variant II (truncation) on the right side.

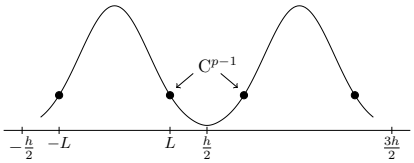


Figure 4: Variant III (regularization).

The main advantage of Variant III is that we are able to construct a function of arbitrary smoothness p while the period h can be chosen relatively small compared to the interval length $2L$. On the other hand, the fact that the approximated functions in Variant I are C^∞ makes this approach spectrally accurate. Using Variant II allows us to choose h relatively small. But, the functions are only continuous and of no higher smoothness. Thus, the Fourier coefficients only decrease rather slow, which also results in the choice of a large cutoff M.

2.2 Nonequispaced discrete Fourier transform (NDFT)

Let the dimension $d \in \mathbb{N}$, the torus $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d \simeq [-1/2, 1/2)^d$ and the sampling set $\mathcal{X} := \{\boldsymbol{x}_j \in \mathbb{T}^d \colon j = 1, \ldots, N\}$ with $N \in \mathbb{N}$ be given. Furthermore, let the multi degree $\boldsymbol{M} = (M_1,\, M_2, \ldots, M_d)^\top \in 2\mathbb{N}^d$ and the index set of possible frequencies

$$\mathcal{I}_{\boldsymbol{M}} := \left\{-\tfrac{M_1}{2}, \ldots, \tfrac{M_1}{2} - 1\right\} \times \ldots \times \left\{-\tfrac{M_d}{2}, \ldots, \tfrac{M_d}{2} - 1\right\}$$

be given. We define the space of d-variate trigonometric polynomials $T_{\boldsymbol{M}}$ of multi degree $\boldsymbol{M}$ by

$$T_{\boldsymbol{M}} := \operatorname{span}\left\{\mathrm{e}^{-2\pi \mathrm{i}\boldsymbol{k}\cdot(\cdot)} \colon \boldsymbol{k} \in I_{\boldsymbol{M}}\right\}.$$

The dimension of this space and, hence, the total number of Fourier coefficients is $|\mathcal{I}_{\boldsymbol{M}}| = M_1 \cdot \ldots \cdot M_d$. Note that we abbreviate the inner product between the frequency $\boldsymbol{k}$ and the time/spatial node $\boldsymbol{x}$ by $\boldsymbol{k} \cdot \boldsymbol{x} = k_1 x_1 + k_2 x_2 + \ldots + k_d x_d$. For clarity of presentation the multi index $\boldsymbol{k}$ addresses elements of vectors and matrices as well.

For a finite number $|\mathcal{I}_{\boldsymbol{M}}|$ of given Fourier coefficients $\hat{f}_{\boldsymbol{k}} \in \mathbb{C},\ \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$, one wants to evaluate the trigonometric polynomial

$$f(\boldsymbol{x}) := \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{f}_{\boldsymbol{k}} \mathrm{e}^{-2\pi \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \in T_{\boldsymbol{M}} \tag{7}$$

at given nonequispaced nodes $\boldsymbol{x}_j \in \mathbb{T}^d$, $j = 1, \ldots, N$. Thus, our concern is the computation of the matrix vector product

$$\boldsymbol{f} = \boldsymbol{A}\hat{\boldsymbol{f}}, \tag{8}$$

where

$$\boldsymbol{f} := \left(f\left(\boldsymbol{x}_j\right)\right)_{j=1,\ldots,N}, \quad \boldsymbol{A} := \left(\mathrm{e}^{-2\pi \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}_j}\right)_{j=1,\ldots,N;\, \boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}, \quad \hat{\boldsymbol{f}} := \left(\hat{f}_{\boldsymbol{k}}\right)_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}.$$

The straightforward algorithm for this matrix vector product, which is called NDFT, takes $\mathcal{O}(N|\mathcal{I}_{\boldsymbol{M}}|)$ arithmetical operations. A related matrix vector product is the adjoint NDFT

$$\hat{\boldsymbol{f}} = \boldsymbol{A}^{\mathsf{H}}\boldsymbol{f}, \qquad \hat{f}_{\boldsymbol{k}} = \sum_{j=1}^{N} f_j \mathrm{e}^{2\pi \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}_j},$$

where $\boldsymbol{A}^{\mathsf{H}} = \overline{\boldsymbol{A}}^{\top}$. Furthermore, note that the inversion formula $\boldsymbol{F}^{-1} = \boldsymbol{F}^{\mathsf{H}}$ for the (equispaced and normalized) Fourier matrix $\boldsymbol{F}$ does **not** hold in the general situation of arbitrary sampling nodes for the matrix $\boldsymbol{A}$.

2.3 Nonequispaced fast Fourier transform (NFFT)

Several algorithms have been proposed for the fast computation of (8), cf.[14,8,49,48,18,21]. In this section we summarize the main ideas of the most successful approach based on[48,32,30]. It makes use an oversampled FFT and a window function φ that is simultaneously well localized in time/space and frequency domain. Given that the window function is well localized in spatial domain, its periodic version

$$\tilde{\varphi}(\boldsymbol{x}) := \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \varphi(\boldsymbol{x} + \boldsymbol{k})$$

is well defined.

Throughout the rest of the tutorial we denote by $\sigma \geq 1$ the oversampling factor and by $m := \sigma M \in \mathbb{N}$ the (oversampled) FFT size (in the case $d = 1$). Furthermore, for $d > 1$ let the vector valued oversampling factor be defined by $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_d)^\top \in \mathbb{R}^d$ (where $\sigma_1, \ldots, \sigma_d \geq 1$) and the FFT size be denoted by $\boldsymbol{m} := \boldsymbol{\sigma} \odot \boldsymbol{M}$. For notational convenience we use the pointwise product

$$\boldsymbol{\sigma} \odot \boldsymbol{M} := (\sigma_1 M_1, \sigma_2 M_2, \ldots, \sigma_d M_d,)^\top$$

and the point wise inverse

$$\boldsymbol{M}^{-1} := \left(M_1^{-1}, M_2^{-1}, \ldots, M_d^{-1}\right)^\top.$$

The main idea is now to approximate the function f by a sum of translates of the one-periodic function $\tilde{\varphi}$, i.e.,

$$f(\boldsymbol{x}) \approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{m}}} g_{\boldsymbol{l}} \tilde{\varphi}(\boldsymbol{x} - \boldsymbol{l} \odot \boldsymbol{m}^{-1}),$$

where we use $\boldsymbol{m} \geq \boldsymbol{M}$ sampling points/translates.

A transformation into Fourier space gives

$$f(\boldsymbol{x}) \approx \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{m}}} \sum_{\boldsymbol{r} \in \mathbb{Z}^d} \hat{g}_{\boldsymbol{k}} c_{\boldsymbol{k}}(\tilde{\varphi}) \mathrm{e}^{-2\pi \mathrm{i}(\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{m}) \cdot \boldsymbol{x}} \tag{9}$$

with the help of the well known convolution theorem. A comparison of (7) and (9) shows that it is reasonable to set

$$\hat{g}_{\boldsymbol{k}} := \begin{cases} \frac{\hat{f}_{\boldsymbol{k}}}{c_{\boldsymbol{k}}(\tilde{\varphi})} & : \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}, \\ 0 & : \text{else.} \end{cases} \tag{10}$$

The final algorithm can basically be divided into three steps (building blocks) and can be summarized as follows.

1. Deconvolve the trigonometric polynomial $f \in T_{\boldsymbol{M}}$ in (7) with a window function in frequency domain, see (10).

2. Compute an FFT on the result of step 1.:

$$g_{\boldsymbol{l}} := \frac{1}{|\mathcal{I}_{\boldsymbol{m}}|} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{m}}} \hat{g}_{\boldsymbol{k}} \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k} \cdot (\boldsymbol{l} \odot \boldsymbol{m}^{-1})}, \quad \boldsymbol{l} \in \mathcal{I}_{\boldsymbol{m}}.$$

3. Convolve the result of step 2. with the window function in time/spatial domain, i.e., evaluate this convolution at the nodes $\boldsymbol{x}_j$, $j = 1, \ldots, N$:

$$f(\boldsymbol{x}_j) \approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{m}}} g_{\boldsymbol{l}} \tilde{\varphi}(\boldsymbol{x}_j - \boldsymbol{l} \odot \boldsymbol{m}^{-1}).$$

Obviously, we end up with a complexity of $\mathcal{O}(|\mathcal{I}_{\boldsymbol{M}}| \log |\mathcal{I}_{\boldsymbol{M}}| + N)$ arithmetic operations. Thereby, the prefactors depend on the required accuracy as well as the properties of the window function. For a description of the NFFT in its matrix-vector notation we refer to[48]. Note that the adjoint NDFT can be approximated in a very similar way by an adjoint NFFT and yields the same enhanced arithmetic complexity. Error bounds in the

∞–norm have already been derived for a variety of possible window functions, see[49,47] for instance. For error estimates in the L_2-norm as well as an automated tuned NFFT see[39]. A widely used implementation is available as part of the NFFT package[29] and is based on the FFTW[20]. This package also offers support of shared memory parallelism[50]. A parallel implementation for graphic processing units was proposed in[31]. Furthermore, an MPI-based parallel NFFT (PNFFT) implementation with support for distributed memory parallelism was proposed in[45] and is publicly available[42]. It is based on a highly scalable MPI extension of FFTW called PFFT[43,41].

3 Fast Ewald Summation for 3d-Periodic Boundary Conditions

For an electrical neutral system (1) of N charges q_j distributed in a cubic box of edge length L we define the electrostatic potential subject to 3d-periodic boundary conditions by

$$\phi^{\mathrm{p3}}(\boldsymbol{x}_j) := \phi_{\mathbb{Z}^3}(\boldsymbol{x}_j) = \sum_{s=0}^{\infty} \sum_{\substack{\boldsymbol{n}\in\mathbb{Z}^3 \\ \|\boldsymbol{n}\|^2=s}} {\sum_{i=1}^{N}}' \frac{q_i}{\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|},$$

i.e., we set $\mathcal{S} := \mathbb{Z}^3$ within the definition (3). Remember that the order of summation has to be specified because of the conditional convergence of the infinite sum, as already pointed out in the introduction.

The following formula was at first presented in[16] by using the Ewald splitting. For a derivation based on convergence factors, see[12]. We have

$$\phi^{\mathrm{p3}}(\boldsymbol{x}_j) = \phi^{\mathrm{p3,S}}(\boldsymbol{x}_j) + \phi^{\mathrm{p3,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p3,self}}(\boldsymbol{x}_j), \tag{11}$$

where for the splitting parameter $\alpha > 0$ we define the short range part

$$\phi^{\mathrm{p3,S}}(\boldsymbol{x}_j) := \sum_{\boldsymbol{n}\in\mathbb{Z}^3} {\sum_{i=1}^{N}}' q_i \frac{\operatorname{erfc}(\alpha\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|},$$

the long range part

$$\phi^{\mathrm{p3,L}}(\boldsymbol{x}_j) := \frac{1}{\pi L} \sum_{\boldsymbol{k}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}} \frac{\mathrm{e}^{-\pi^2\|\boldsymbol{k}\|^2/(\alpha^2L^2)}}{\|\boldsymbol{k}\|^2} \left(\sum_{i=1}^{N} q_i \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}_i/L} \right) \mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}_j/L},$$

and the self potential

$$\phi^{\mathrm{p3,self}}(\boldsymbol{x}_j) := -\frac{2\alpha}{\sqrt{\pi}} q_j.$$

Often a fourth term, the so called dipole correction, appears in the decomposition (11), cf.[13]. The dipole correction term is the only part depending on the order of summation. However, if a spherical summation order is applied, the dipole correction term depends only on the norm of the dipole moment $\sum_{j=1}^{N} q_j\boldsymbol{x}_j$ and, additionally, on the dielectric constant of the surrounding medium. Therefore, it can be computed efficiently in $\mathcal{O}(N)$ arithmetic operations. If the medium is assumed to be metallic, the dipole term vanishes

and (11) applies. It should be mentioned that the formulas above can be generalized to non-cubic boxes and also non-orthogonal (triclinic) boxes, cf.[16,11,27].

Since the complementary error function erfc rapidly tends to zero, the short range part of each potential $\phi^{\mathrm{p3,S}}(\boldsymbol{x}_j)$ can be obtained by direct evaluation, i.e., all distances $\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|$ larger than an appropriate cutoff radius $r_{\mathrm{cut}} > 0$ are ignored. If we assume a sufficiently homogenous particle distribution, each particle only interacts with a fixed number of neighbors. Thus, the real space sum can be computed with a linked cell algorithm[19] in $\mathcal{O}(N)$ arithmetic operations for this case.

In order to compute the long range parts $\phi^{\mathrm{p3,L}}(\boldsymbol{x}_j)$ we truncate the infinite sum and compute approximations of the sums

$$\hat{S}(\boldsymbol{k}) := \sum_{i=1}^{N} q_i \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}\cdot\boldsymbol{x}_i/L}, \quad \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}},$$

with an adjoint NFFT and evaluate

$$\phi^{\mathrm{p3,L}}(\boldsymbol{x}_j) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}\setminus\{\boldsymbol{0}\}} \hat{b}_{\boldsymbol{k}} \hat{S}(\boldsymbol{k}) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}\cdot\boldsymbol{x}_j/L}, \quad j = 1,\dots,N,$$

via the NFFT. Thereby, we define the Fourier coefficients

$$\hat{b}_{\boldsymbol{k}} := \frac{1}{\pi L} \frac{\mathrm{e}^{-\pi^2\|\boldsymbol{k}\|^2/(\alpha^2 L^2)}}{\|\boldsymbol{k}\|^2}.$$

The proposed evaluation of $\phi^{\mathrm{p3,L}}(\boldsymbol{x}_j)$ at the points $\boldsymbol{x}_j$, $j = 1,\dots,N$, requires $\mathcal{O}(N + |\mathcal{I}_{\boldsymbol{M}}| \log |\mathcal{I}_{\boldsymbol{M}}|)$ arithmetic operations.

In matrix vector notation we may write

$$\left(\phi^{\mathrm{p3,L}}(\boldsymbol{x}_j)\right)_{j=1}^{N} \approx \tilde{\boldsymbol{A}}\boldsymbol{D}\tilde{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{q}, \tag{12}$$

where $\tilde{\boldsymbol{A}} \approx \boldsymbol{A}$ denotes the matrix representation of the NFFT (≈NDFT) in three dimensions, $\boldsymbol{D}$ is a diagonal matrix with entries $\hat{b}_{\boldsymbol{k}}$, $\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$, and $\boldsymbol{q} = (q_1,\dots,q_N)^{\mathrm{T}} \in \mathbb{R}^N$.

Relations to existing work

A straightforward method, that accelerates the traditional Ewald summation technique by NFFT was already presented in[44]. This combination was first presented in[25] is very similar to the FFT-accelerated Ewald sum methods, namely, the so-called particle-particle particle-mesh ($\mathrm{P^3M}$), particle-mesh Ewald (PME) and smooth particle-mesh Ewald (SPME), see[13] and also[51].

4 Fast Ewald Summation for 2d-Periodic Boundary Conditions

In this section we denote for $\boldsymbol{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ the vector of its first two components by $\tilde{\boldsymbol{y}} := (y_1, y_2) \in \mathbb{R}^2$, $j = 1,\dots,N$. We consider an electrical neutral system (1) of N charges $q_j \in \mathbb{R}$ at positions $\boldsymbol{x}_j = (\tilde{\boldsymbol{x}}_j, x_{j,3}) \in L\mathbb{T}^2 \times \mathbb{R}$. Under periodic boundary conditions in the first two dimensions we define the potential of each single particle by

$$\phi^{\mathrm{p2}}(\boldsymbol{x}_j) := \phi_{\mathbb{Z}^2\times\{0\}}(\boldsymbol{x}_j) = \sum_{s=0}^{\infty} \sum_{\substack{\boldsymbol{n}\in\mathbb{Z}^2\times\{0\} \\ \|\boldsymbol{n}\|^2=s}} {\sum_{i=1}^{N}}' \frac{q_i}{\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|},$$

i.e., we set $\mathcal{S} := \mathbb{Z}^2 \times \{0\}$ within the definition (3). This can be rewritten in the form

$$\phi^{\mathrm{p2}}(\boldsymbol{x}_j) = \phi^{\mathrm{p2,S}}(\boldsymbol{x}_j) + \phi^{\mathrm{p2,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p2},\mathbf{0}}(\boldsymbol{x}_j) + \phi^{\mathrm{p2,self}}(\boldsymbol{x}_j), \tag{13}$$

where for some $\alpha > 0$ we define the short range part

$$\phi^{\mathrm{p2,S}}(\boldsymbol{x}_j) := \sum_{\boldsymbol{n}\in\mathbb{Z}^2\times\{0\}} \sum_{i=1}^{N}{}' q_i \frac{\operatorname{erfc}(\alpha\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|},$$

the long range parts

$$\begin{aligned}
\phi^{\mathrm{p2,L}}(\boldsymbol{x}_j) &:= \frac{1}{2L} \sum_{\boldsymbol{k}\in\mathbb{Z}^2\setminus\{\mathbf{0}\}} \sum_{i=1}^{N} q_i \, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}\cdot\tilde{\boldsymbol{x}}_{ij}/L} \cdot \Theta^{\mathrm{p2}}(\|\boldsymbol{k}\|, x_{ij,3}), \\
\phi^{\mathrm{p2},\mathbf{0}}(\boldsymbol{x}_j) &:= -\frac{2\sqrt{\pi}}{L^2} \sum_{i=1}^{N} q_i \Theta_0^{\mathrm{p2}}(x_{ij,3}),
\end{aligned} \tag{14}$$

the self potential

$$\phi^{\mathrm{p2,self}}(\boldsymbol{x}_j) := -\frac{2\alpha}{\sqrt{\pi}} q_j \ ,$$

and the functions $\Theta^{\mathrm{p2}}(k, r)$, $\Theta_0^{\mathrm{p2}}(r)$ for $k, r \in \mathbb{R}$ are defined by

$$\begin{aligned}
\Theta^{\mathrm{p2}}(k, r) &:= \frac{1}{k}\left[\mathrm{e}^{2\pi k r/L} \operatorname{erfc}\left(\frac{\pi k}{\alpha L} + \alpha r\right) + \mathrm{e}^{-2\pi k r/L} \operatorname{erfc}\left(\frac{\pi k}{\alpha L} - \alpha r\right)\right], \\
\Theta_0^{\mathrm{p2}}(r) &:= \frac{\mathrm{e}^{-\alpha^2 r^2}}{\alpha} + \sqrt{\pi} r \operatorname{erf}(\alpha r).
\end{aligned}$$

These expressions were already given in[23]. In the Appendix of[40] we give a proof using convergence factors, similar to the proof of the 3d-periodic case in[12]. Thereby, we always start with the splitting (4) and then use the technique of convergence factors to derive the Fourier space representation of the long range part by applying the Poisson summation formula.

The evaluation of the short range part $\phi^{\mathrm{p2,S}}(\boldsymbol{x}_j)$ is again done by a direct evaluation. For the computation of the long range part we truncate the infinite sum in $\phi^{\mathrm{p2,L}}(\boldsymbol{x}_j)$, i.e., for some appropriate $\tilde{\boldsymbol{M}} = (M_1, M_2) \in 2\mathbb{N}^2$ we set

$$\phi^{\mathrm{p2,L}}(\boldsymbol{x}_j) \approx \frac{1}{2L} \sum_{\boldsymbol{k}\in\mathcal{I}_{\tilde{\boldsymbol{M}}}\setminus\{\mathbf{0}\}} \sum_{i=1}^{N} q_i \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}\cdot\tilde{\boldsymbol{x}}_{ij}/L} \Theta^{\mathrm{p2}}(\|\boldsymbol{k}\|, x_{ij,3}).$$

and apply the regularization Variant III from Section 2.1 to the functions $\Theta^{\mathrm{p2}}(\|\boldsymbol{k}\|, \cdot)$. To this end we assume without loss of generality $L_3 > 0$ large enough such that $x_{j,3} \in [-L_3/2, L_3/2]$, i.e., the particle coordinates are bounded also in the non-periodic dimension. Thus, all the functions $\Theta^{\mathrm{p2}}(\|\boldsymbol{k}\|, \cdot)$ have to be evaluated only within the finite interval $[-L_3, L_3]$. Note that we have to double the interval length since we do not have periodicity in the last dimension. The same approximation idea is applied to the kernel function $\Theta_0^{\mathrm{p2}}(r)$ in (14). Note that $\lim_{x\to\pm\infty}[\mathrm{e}^{-x^2} + \sqrt{\pi}x \ \operatorname{erf}(x)] = \lim_{x\to\pm\infty} |x| = \infty$, i.e., the approximation Variant I given in Section 2.1 is not applicable.

At first, we choose $h > 2L_3$ and accordingly some $\varepsilon \in (0, 1/2)$ such that $|x_{ij,3}| \leq L_3 =: h(1/2 - \varepsilon) < h/2$ for all $i, j = 1, \ldots, N$. This corresponds to a surrounding box that is large enough to hold all differences of particle coordinates in the last dimension. In addition, since the strong inequality $h > 2L_3$ holds we have some extra space for constructing a regularization. In order to approximate the long range parts $\phi^{\mathrm{p2,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p2},\mathbf{0}}(\boldsymbol{x}_j)$ efficiently we consider for $k \in \{\|\boldsymbol{k}\| : \boldsymbol{k} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}\}$ the regularizations

$$K_{\mathrm{R}}(k,r) := \begin{cases} \dfrac{1}{2L}\Theta^{\mathrm{p2}}(k,r) & : k \neq 0, |h^{-1}r| \leq 1/2 - \varepsilon, \\ -\dfrac{2\sqrt{\pi}}{L^2}\Theta_0^{\mathrm{p2}}(r) & : k = 0, |h^{-1}r| \leq 1/2 - \varepsilon, \\ K_{\mathrm{B}}(k,r) & : |h^{-1}r| \in (1/2 - \varepsilon, 1/2], \end{cases} \tag{15}$$

where we claim that each function $K_{\mathrm{B}}(k,\cdot) : [-h/2, -h/2 + h\varepsilon] \cup [h/2 - h\varepsilon, h/2] \to \mathbb{R}$ fulfills the Hermite interpolation conditions

$$\frac{\partial^j}{\partial r^j} K_{\mathrm{B}}(k, h/2 - h\varepsilon) = \begin{cases} \frac{1}{2L}\frac{\partial^j}{\partial r^j}\Theta^{\mathrm{p2}}(k, h/2 - h\varepsilon) & : k \neq 0, \\ -\frac{2\sqrt{\pi}}{L^2}\frac{\mathrm{d}^j}{\mathrm{d}r^j}\Theta_0^{\mathrm{p2}}(h/2 - h\varepsilon) & : k = 0, \end{cases} \tag{16}$$

$$\frac{\partial^j}{\partial r^j} K_{\mathrm{B}}(k, -h/2 + h\varepsilon) = \begin{cases} \frac{1}{2L}\frac{\partial^j}{\partial r^j}\Theta^{\mathrm{p2}}(k, -h/2 + h\varepsilon) & : k \neq 0, \\ -\frac{2\sqrt{\pi}}{L^2}\frac{\mathrm{d}^j}{\mathrm{d}r^j}\Theta_0^{\mathrm{p2}}(-h/2 + h\varepsilon) & : k = 0, \end{cases} \tag{17}$$

for all $j = 0, \ldots, p-1$. Hereby, we refer to $p \in \mathbb{N}$ as the degree of smoothness. In order to end up with h-periodic, smooth functions $K_{\mathrm{R}}(k,\cdot)$, the functions $K_{\mathrm{B}}(k,\cdot)$ are constructed such that

$$\frac{\partial^j}{\partial r^j} K_{\mathrm{R}}(k, h/2) = \frac{\partial^j}{\partial r^j} K_{\mathrm{R}}(k, -h/2) \quad \text{for } j = 0, \ldots, p-1$$

is also fulfilled. In Theorem 2.1 we show that the functions $K_{\mathrm{B}}(k,.)$ can be constructed as polynomials of degree $2p-1$ by two point Taylor interpolation. Figure 5 shows an example of such a regularization $K_{\mathrm{R}}(k,\cdot)$.

In summary, the functions $K_{\mathrm{R}}(k,\cdot)$ are h-periodic and smooth, i.e., $K_{\mathrm{R}}(k,\cdot) \in \mathrm{C}^{p-1}(h\mathbb{T})$. Therefore, they can be approximated by a truncated Fourier series up to a prescribed error. To this end, we approximate for each $k \in \{\|\boldsymbol{k}\| \neq 0 : \boldsymbol{k} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}\}$ the function

$$\frac{1}{2L}\Theta^{\mathrm{p2}}(k,r) \approx \sum_{l \in \mathcal{I}_{M_3}} \hat{b}_{k,l} \mathrm{e}^{2\pi \mathrm{i} l r/h} \tag{18}$$

for $|r| \leq h/2 - h\varepsilon = L_3$ by the truncated Fourier series of its regularization $K_{\mathrm{R}}(k,\cdot)$. Analogously, for $k = 0$ we have

$$-\frac{2\sqrt{\pi}}{L^2}\Theta_0^{\mathrm{p2}}(r) \approx \sum_{l \in \mathcal{I}_{M_3}} \hat{b}_{0,l} \mathrm{e}^{2\pi \mathrm{i} l r/h}. \tag{19}$$

Thereby, we choose the frequency cutoff $M_3 \in 2\mathbb{N}$ large enough and compute the Fourier coefficients $\hat{b}_{k,l}$ in (18) as well as $\hat{b}_{0,l}$ in (19) by the discrete Fourier transform

$$\hat{b}_{k,l} := \frac{1}{M_3} \sum_{j \in \mathcal{I}_{M_3}} K_{\mathrm{R}}\left(k, \tfrac{jh}{M_3}\right) \mathrm{e}^{-2\pi \mathrm{i} j l/M_3}, \qquad l = -M_3/2, \ldots, M_3/2 - 1.$$

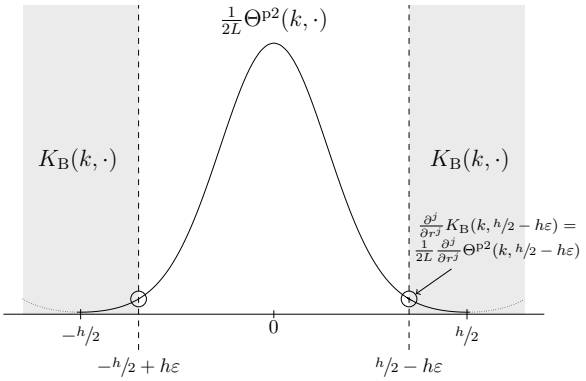


Figure 5: Example for $K_{\mathrm{R}}(k,\cdot)$ for $k \geq 1$. At the boundaries (gray area) the regularization adopts the values of the boundary function $K_{\mathrm{B}}(k,\cdot)$. We also marked the points, where the conditions (16) and (17) are fulfilled. In our implementation, the function in the gray area is a polynomial of degree $2p-1$ constructed by two-point Taylor interpolation.

This ansatz is closely related to the fast summation method described in[47]. Due to the fact that we have $\Theta_0^{\mathrm{p}2}(\cdot), \Theta^{\mathrm{p}2}(k,\cdot) \in \mathrm{C}^\infty(\mathbb{R})$ $(k \geq 1)$ we are not restricted in the choice of the parameter p. By choosing M_3 large enough we can construct approximations (18) and (19) of a required accuracy.

If $k \in \{\|\boldsymbol{k}\| \neq 0 : \boldsymbol{k} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}\}$ is large enough, then the function value $\Theta^{\mathrm{p}2}(k, h/2)$ might be sufficiently small so that

$$\Theta^{\mathrm{p}2}(k,r) \approx \sum_{n\in\mathbb{Z}} \Theta^{\mathrm{p}2}(k, r+hn),$$

yields a good approximation, see Figure 6.

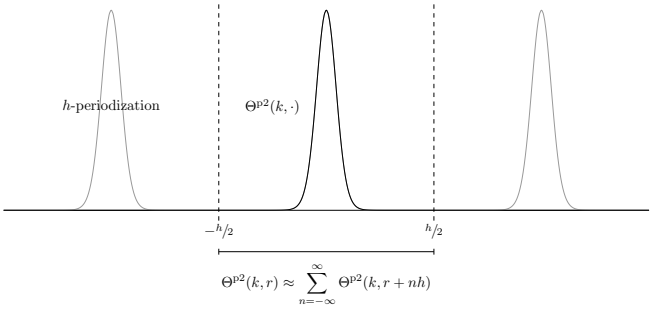


Figure 6: If k is sufficiently large, the h-periodic version of $\Theta^{\mathrm{p}2}(k,\cdot)$ might be a good approximation of $\Theta^{\mathrm{p}2}(k,\cdot)$.

In this case we could also apply Variant I, as described in Section 2. The analytical

Fourier transform of $\Theta^{\mathrm{p2}}(k,\cdot)$ is given by

$$\hat{\Theta}^{\mathrm{p2}}(k,\xi) = \int_{-\infty}^{\infty} \Theta^{\mathrm{p2}}(k,r)\mathrm{e}^{-2\pi \mathrm{i} r\xi}\mathrm{d}r = \frac{2L}{\pi(k^2+L^2\xi^2)}\mathrm{e}^{-\pi^2k^2/(\alpha^2L^2)-\pi^2\xi^2/\alpha^2},$$

see[34], for instance. Applying the Poisson summation formula leads to

$$\frac{1}{2L}\Theta^{\mathrm{p2}}(k,r) \approx \frac{1}{2L}\sum_{n\in\mathbb{Z}} \Theta^{\mathrm{p2}}(k,r+hn) \approx \frac{1}{2Lh}\sum_{l\in\mathcal{I}_{M_3}} \hat{\Theta}^{\mathrm{p2}}(k,{}^{l}\!/_{h})\mathrm{e}^{2\pi \mathrm{i} lr/h},$$

i.e., we can simply set $\hat{b}_{k,l} := (2Lh)^{-1}\hat{\Theta}^{\mathrm{p2}}(k,{}^{l}\!/_{h})$ instead of regularizing the function.

In summary, we obtain the following approximation for the long range parts,

$$\begin{aligned}\phi^{\mathrm{p2,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p2},\mathbf{0}}(\boldsymbol{x}_j) &\approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\tilde{\boldsymbol{M}}}}\sum_{l\in\mathcal{I}_{M_3}} \hat{b}_{\|\boldsymbol{k}\|,l}\sum_{i=1}^{N} q_i\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}\cdot\tilde{\boldsymbol{x}}_{ij}/L}\mathrm{e}^{2\pi\mathrm{i}lx_{ij,3}/h}\\ &= \sum_{(\boldsymbol{k},l)\in\mathcal{I}_{\boldsymbol{M}}} \hat{b}_{\|\boldsymbol{k}\|,l}\left(\sum_{i=1}^{N} q_i\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{v}(\boldsymbol{k},l)\cdot\boldsymbol{x}_i}\right)\mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{v}(\boldsymbol{k},l)\cdot\boldsymbol{x}_j},\end{aligned}$$

where we substitute the truncated Fourier series (18), (19) into (13), (14) and define $\boldsymbol{M} := (\tilde{\boldsymbol{M}}, M_3) \in 2\mathbb{N}^3$ as well as the vectors $\boldsymbol{v}(\boldsymbol{k},l) := (\boldsymbol{k}/L, l/h) \in L^{-1}\mathbb{Z}^2 \times h^{-1}\mathbb{Z}$. The expressions in the inner brackets

$$\hat{S}(\boldsymbol{k},l) := \sum_{i=1}^{N} q_i\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{v}(\boldsymbol{k},l)\cdot\boldsymbol{x}_i}, \quad (\boldsymbol{k},l)\in\mathcal{I}_{\boldsymbol{M}},$$

can be computed by an adjoint NFFT. This will be followed by $|\mathcal{I}_{\boldsymbol{M}}|$ multiplications with $\hat{b}_{\|\boldsymbol{k}\|,l}$ and completed by an NFFT to compute the outer summation over the indexes $(\boldsymbol{k},l)\in\mathcal{I}_{\boldsymbol{M}}$. Therefore, the proposed evaluation of $\phi^{\mathrm{p2,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p2},\mathbf{0}}(\boldsymbol{x}_j)$ at the points $\boldsymbol{x}_j$, $j=1,\dots,N$, requires $\mathcal{O}(N + |\mathcal{I}_{\boldsymbol{M}}|\log|\mathcal{I}_{\boldsymbol{M}}|)$ arithmetic operations.

We obtain a similar matrix-vector notation as in the 3d-periodic case, namely,

$$\left(\phi^{\mathrm{p2,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p2},0}(\boldsymbol{x}_j)\right)_{j=1}^{N} \approx \tilde{\boldsymbol{A}}\boldsymbol{D}\tilde{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{q}, \tag{20}$$

where $\tilde{\boldsymbol{A}}$ denotes the matrix representation of the NFFT in three dimensions for the nodes $(\tilde{\boldsymbol{x}}_j/L, x_{j,3}/h) \in \mathbb{T}^3$, $\boldsymbol{D}$ is a diagonal matrix with entries $\hat{b}_{\|\boldsymbol{k}\|,l}$, $(\boldsymbol{k},l)\in\mathcal{I}_{\boldsymbol{M}}$, and $\boldsymbol{q} = (q_1,\dots,q_N)^{\top} \in \mathbb{R}^N$.

Relations to existing work

The Ewald formulas (13) for 2d-periodic geometries were already proposed in[23]. We remark that a method based on the splitting (13) is used in[34] in combination with Variant I of Section 2. As pointed out on page 12 in[34] this approach is limited to functions that decay sufficiently fast in the interval $[-{}^{h}\!/_{2}, {}^{h}\!/_{2})$. In other words, whenever $\Theta^{\mathrm{p2}}(k,\max|x_{ij,3}|)$ is not sufficiently small we need to choose a relatively large period $h \gg 2L$, which may also result in the choice of a large cutoff M_3. Some other Fourier based algorithms, like MMM2D[6] or ELC[4] already exist. A method based on approximation Variant II from Section 2.1 is proposed in[38]. However, as mentioned in Section 2.1 this method suffers from the rather slow convergence rate in Fourier space. See also[52,9] for algorithms with higher complexity.

5 Fast Ewald Summation for 1d-Periodic Boundary Conditions

In this section we denote for some $\boldsymbol{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ the vector of its last two components by $\tilde{\boldsymbol{y}} := (y_2, y_3) \in \mathbb{R}^2$. We consider a system of N charges $q_j \in \mathbb{R}$ at positions $\boldsymbol{x}_j = (x_{j,1}, \tilde{\boldsymbol{x}}_j) \in L\mathbb{T} \times \mathbb{R}^2$, $j = 1, \ldots, N$. If periodic boundary conditions are assumed only in the first coordinate we define the potential of each single particle j by

$$\phi^{\mathrm{p1}}(\boldsymbol{x}_j) := \phi_{\mathbb{Z}\times\{0\}^2}(\boldsymbol{x}_j) = \sum_{s=0}^{\infty} \sum_{\substack{\boldsymbol{n}\in\mathbb{Z}\times\{0\}^2 \\ |n_1|=s}} \sum_{i=1}^{N}{}' \frac{q_i}{\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|} \tag{21}$$

i.e., we set $\mathcal{S} := \mathbb{Z} \times \{0\}^2$ within definition (3). In the following we denote by

$$\Gamma(s, x) := \int_x^{\infty} t^{s-1} \mathrm{e}^{-t} \mathrm{d}t$$

the upper incomplete gamma function. For the case $s = 0$ the well known identity

$$\Gamma(0, x) = -\gamma - \ln x - \sum_{k=1}^{\infty} (-1)^k \frac{x^k}{k!k}$$

holds for all positive x, see [number 5.1.11] in[2]. Thereby, γ is the Euler-Mascheroni constant. The function $\Gamma(0, \cdot)$ is also known as the exponential integral function. We easily see

$$\lim_{x\to 0} \Gamma(0, x) + \ln x + \gamma = 0.$$

The potential (21) can be written as

$$\phi^{\mathrm{p1}}(\boldsymbol{x}_j) = \phi^{\mathrm{p1,S}}(\boldsymbol{x}_j) + \phi^{\mathrm{p1,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p1,0}}(\boldsymbol{x}_j) + \phi^{\mathrm{p1,self}}(\boldsymbol{x}_j),$$

where for the splitting parameter $\alpha > 0$ we define the short range part

$$\phi^{\mathrm{p1,S}}(\boldsymbol{x}_j) := \sum_{\boldsymbol{n}\in\mathbb{Z}\times\{0\}^2} \sum_{i=1}^{N}{}' q_i \frac{\mathrm{erfc}(\alpha\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij} + L\boldsymbol{n}\|},$$

the long range parts

$$\phi^{\mathrm{p1,L}}(\boldsymbol{x}_j) := \frac{2}{L} \sum_{k\in\mathbb{Z}\setminus\{0\}} \sum_{i=1}^{N} q_i \, \mathrm{e}^{2\pi \mathrm{i} k(x_{i,1} - x_{j,1})/L} \cdot \Theta^{\mathrm{p1}}(k, \|\tilde{\boldsymbol{x}}_{ij}\|)\,, \tag{22}$$

$$\phi^{\mathrm{p1,0}}(\boldsymbol{x}_j) := -\frac{1}{L} \sum_{\substack{i=1 \\ \|\tilde{\boldsymbol{x}}_{ij}\|\neq 0}}^{N} q_i \Theta_0^{\mathrm{p1}}(\|\tilde{\boldsymbol{x}}_{ij}\|), \tag{23}$$

the self potential

$$\phi^{\mathrm{p1,self}}(\boldsymbol{x}_j) := -\frac{2\alpha}{\sqrt{\pi}} q_j\,,$$

and the functions $\Theta^{\mathrm{p1}}(k,r), \Theta_0^{\mathrm{p1}}(r)$ for $k, r \in \mathbb{R}$ are defined by

$$\Theta^{\mathrm{p1}}(k,r) := \int_0^\alpha \frac{1}{z} \mathrm{e}^{\frac{-\pi^2 k^2}{L^2 z^2}} \mathrm{e}^{-r^2 z^2} \mathrm{d}z,$$
$$\Theta_0^{\mathrm{p1}}(k,r) := \gamma + \Gamma(0, \alpha^2 r^2) + \ln(\alpha^2 r^2)\,.$$

The function $\Theta^{\mathrm{p1}}(k,r)$ can be expressed by the incomplete modified Bessel function of the second kind[24], see Section 5.2.2 in[40]. This function is known to be indefinitely often differentiable and, thus, we can construct regularizations of similar structure as (15) in order to construct a fast algorithm. In this case the final algorithm requires a smooth bivariate regularization, which can be obtained easily from a one dimensional construction as the Fourier coefficients are radial in $\tilde{\boldsymbol{x}}_{ij}$.

By the Lemma 5.2 in[40] we show that the function $\Theta^{\mathrm{p1}}(k,r)$ for fixed r tends to zero exponentially fast for growing k, which allows the truncation of the infinite sum in $\phi^{\mathrm{p1,L}}(\boldsymbol{x}_j)$. Furthermore, Lemma 5.3 in[40] shows that also the kernel in $\phi^{\mathrm{p1,0}}(\boldsymbol{x}_j)$ is a smooth function, which allows the application of the fast summation method. Note that we have $\lim_{x\to\pm\infty} \gamma + \Gamma(0,x^2) + \ln(x^2) = \infty$. Thus, the approximation Variant I given in Section 2.1 is not applicable, just as in the case of the $\boldsymbol{k} = \mathbf{0}$ term of the 2d-periodic Ewald sum. However, using the fast summation approach, the function is truncated and embedded in a smooth and periodic function, which does not require localization of the kernel function.

Similar as in the previous section we derive the fast algorithm based on (22) and (23). The evaluation of the short range part $\phi^{\mathrm{p1,S}}(\boldsymbol{x}_j)$ is done by a direct evaluation again. Due to Lemma 5.2 in[40] we truncate the infinite sum in $\phi^{\mathrm{p1,L}}(\boldsymbol{x}_j)$, i.e., for some appropriate $M_1 \in 2\mathbb{N}$ we set

$$\phi^{\mathrm{p1,L}}(\boldsymbol{x}_j) \approx \frac{2}{L} \sum_{k \in \mathcal{I}_{M_1} \setminus \{0\}} \sum_{i=1}^{N} q_i \mathrm{e}^{2\pi \mathrm{i} k x_{ij,1}/L} \Theta^{\mathrm{p1}}(k, \|\tilde{\boldsymbol{x}}_{ij}\|).$$

In the following we assume that $\tilde{\boldsymbol{x}}_j \in [-L_2/2, L_2/2] \times [-L_3/2, L_3/2]$, i.e., $\tilde{\boldsymbol{x}}_{ij} \in [-L_2, L_2] \times [-L_3, L_3]$. Thus, the particle distances regarding the non-periodic dimensions $\|\tilde{\boldsymbol{x}}_{ij}\|$ are bounded above by $\sqrt{L_2^2 + L_3^2}$. Furthermore, we choose some $h > 2\sqrt{L_2^2 + L_3^2}$ and accordingly some $\varepsilon \in (0, 1/2)$ such that $\|\tilde{\boldsymbol{x}}_{ij}\| \leq \sqrt{L_2^2 + L_3^2} =: h(1/2 - \varepsilon) < h/2$ for all $i, j = 1, \ldots, N$.

In order to approximate the long range part $\phi^{\mathrm{p1,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p1,0}}(\boldsymbol{x}_j)$ efficiently we consider for $k \in \{0, \ldots, M_1/2\}$ the regularizations

$$K_{\mathrm{R}}(k,r) := \begin{cases} \frac{2}{L}\Theta^{\mathrm{p1}}(k,r) & : k \neq 0, |h^{-1}r| \leq 1/2 - \varepsilon, \\ -\frac{1}{L}\Theta_0^{\mathrm{p1}}(r) & : k = 0, |h^{-1}r| \leq 1/2 - \varepsilon, \\ K_{\mathrm{B}}(k,r) & : |h^{-1}r| \in (1/2 - \varepsilon, 1/2], \\ K_{\mathrm{B}}(k, h/2) & : |h^{-1}r| > 1/2, \end{cases},$$

where each function $K_{\mathrm{B}}(k,\cdot) : [h/2 - h\varepsilon, h/2] \to \mathbb{R}$ is constructed such that $K_{\mathrm{R}}(k, \|\cdot\|) : h\mathbb{T}^2 \to \mathbb{R}$ is in the Sobolev space $\mathrm{C}^{p-1}(h\mathbb{T}^2)$, i.e., $K_{\mathrm{B}}(k,\cdot)$ fulfills the interpolation

conditions

$$\frac{\partial^j}{\partial r^j} K_{\mathrm{B}}(k, {}^h\!/_2 - h\varepsilon) = \begin{cases} \frac{2}{L} \frac{\partial^j}{\partial r^j} \Theta^{\mathrm{p1}}(k, {}^h\!/_2 - h\varepsilon) & : k \neq 0, \\ -\frac{1}{L} \frac{\mathrm{d}^j}{\mathrm{d}r^j} \Theta_0^{\mathrm{p1}}({}^h\!/_2 - h\varepsilon) & : k = 0 \end{cases} \tag{24}$$

for $j = 0, \ldots, p-1$ as well as

$$\frac{\partial^j}{\partial r^j} K_{\mathrm{B}}(k, {}^h\!/_2) = 0 \quad \text{for } j = 1, \ldots, p-1. \tag{25}$$

Note that $K_{\mathrm{R}}(k, \|\cdot\|)$ is constant for all the points $\{\boldsymbol{y} \in h\mathbb{T}^2 : \|\boldsymbol{y}\| \geq {}^h\!/_2\}$. Therefore, the conditions (25) ensure smoothness of $K_{\mathrm{R}}(k, \|\cdot\|)$ in the points $\{\boldsymbol{y} \in h\mathbb{T}^2 : \|\boldsymbol{y}\| = {}^h\!/_2\}$. Furthermore, (25) does not include any restriction on the function value of $K_{\mathrm{R}}(k, {}^h\!/_2)$, since it does not influence the smoothness of $K_{\mathrm{R}}(k, \|\cdot\|)$. In Appendix C of[40] we show that an adopted version of Theorem 2.1 can be used to construct the regularizing functions $K_{\mathrm{B}}(k, \|.\|)$ as interpolation polynomials of degree $2p-2$. By our construction the functions $K_{\mathrm{R}}(k, \|\cdot\|)$ are h-periodic in each direction and smooth, i.e., $K_{\mathrm{R}}(k, \|\cdot\|) \in \mathrm{C}^{p-1}(h\mathbb{T}^2)$. For a graphical illustration of a regularization $K_{\mathrm{R}}(k, \cdot)$ see Figure 7.

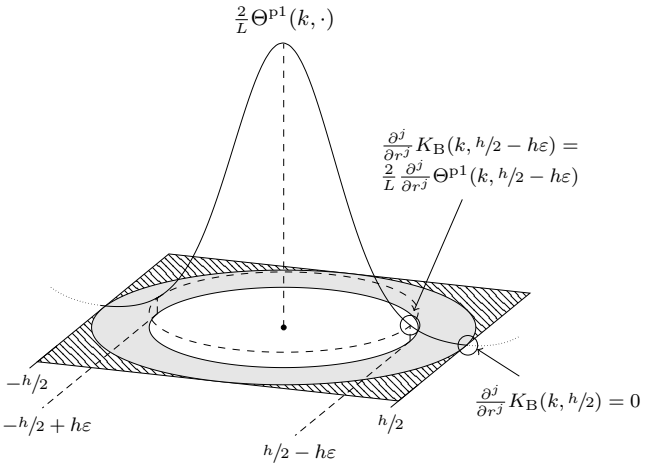


Figure 7: Example for $K_{\mathrm{R}}(k, \cdot)$ for $k \geq 1$. Over the gray area the regularization adopts the values of the boundary function $K_{\mathrm{B}}(k, \cdot)$, which we compute via a modified two point Taylor interpolation. In the corners (striped area) $K_{\mathrm{R}}(k, \cdot)$ has the constant value $K_{\mathrm{B}}(k, {}^h\!/_2)$. We also marked the points, where the conditions (24) and (25) are fulfilled.

To this end, we approximate for each $k \in \mathcal{I}_{M_1} \setminus \{0\}$ the function

$$\frac{2}{L} \Theta^{\mathrm{p1}}(k, \|\boldsymbol{y}\|) \approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\tilde{M}}} \hat{b}_{k,\boldsymbol{l}} \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l} \cdot \boldsymbol{y}/h} \tag{26}$$

for $\|\boldsymbol{y}\| \leq {}^h\!/_2 - h\varepsilon$ by a trigonometric polynomial. In the case $k = 0$ we use the approximation

$$-\frac{1}{L} \Theta_0^{\mathrm{p1}}(\alpha^2 \|\boldsymbol{y}\|^2) \approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\tilde{M}}} \hat{b}_{0,\boldsymbol{l}} \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l} \cdot \boldsymbol{y}/h}. \tag{27}$$

Thereby, we choose $\tilde{\boldsymbol{M}} = (M_2, M_3) \in 2\mathbb{N}^2$ large enough and compute the Fourier coefficients $\hat{b}_{k,\boldsymbol{l}}$ by

$$\hat{b}_{k,\boldsymbol{l}} := \frac{1}{|\mathcal{I}_{\tilde{\boldsymbol{M}}}|} \sum_{\boldsymbol{j} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}} K_{\mathrm{R}}\left(k, \|\boldsymbol{j} \odot \tilde{\boldsymbol{M}}^{-1}\| h\right) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{j} \cdot (\boldsymbol{l} \odot \tilde{\boldsymbol{M}}^{-1})}$$

for all $k \in \mathcal{I}_{M_1}$.

For relatively large values of k we may again obtain a good approximation by setting

$$\Theta^{\mathrm{p1}}(k, \|\boldsymbol{y}\|) \approx \sum_{\boldsymbol{n} \in \mathbb{Z}^2} \Theta^{\mathrm{p1}}(k, \|\boldsymbol{y} + h\boldsymbol{n}\|),$$

compare to the 2d-periodic case and Figure 6. With the help of the analytical Fourier transform

$$\hat{\Theta}^{\mathrm{p1}}(k, \|\boldsymbol{\xi}\|) = \frac{L^2}{2\pi(k^2 + L^2\|\boldsymbol{\xi}\|^2)} \mathrm{e}^{-\pi^2 k^2/(\alpha^2 L^2) - \pi^2 \|\boldsymbol{\xi}\|^2/\alpha^2}$$

and the Poisson summation formula we get

$$\frac{2}{L}\hat{\Theta}^{\mathrm{p1}}(k, \|\boldsymbol{y}\|) \approx \frac{2}{L} \sum_{\boldsymbol{n} \in \mathbb{Z}^2} \Theta^{\mathrm{p1}}(k, \|\boldsymbol{y} + h\boldsymbol{n}\|) \approx \frac{2}{Lh^2} \sum_{\boldsymbol{l} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}} \hat{\Theta}^{\mathrm{p1}}(k, h^{-1}\|\boldsymbol{l}\|) \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l} \cdot \boldsymbol{y}/h},$$

i.e., we can simply set $\hat{b}_{k,\boldsymbol{l}} := 2(Lh^2)^{-1}\hat{\Theta}^{\mathrm{p1}}(k, h^{-1}\|\boldsymbol{l}\|)$ instead of regularizing the function.

In summary we obtain the following approximation for the long range parts

$$\begin{aligned}
\phi^{\mathrm{p1,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p1},0}(\boldsymbol{x}_j) &\approx \sum_{k \in \mathcal{I}_{M_1}} \sum_{\boldsymbol{l} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}} \hat{b}_{|k|,\boldsymbol{l}} \sum_{i=1}^{N} q_i \mathrm{e}^{2\pi \mathrm{i} k x_{ij,1}/L} \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l} \cdot \tilde{\boldsymbol{x}}_{ij}/h} \\
&= \sum_{(k,\boldsymbol{l}) \in \mathcal{I}_{\mathcal{M}}} \hat{b}_{|k|,\boldsymbol{l}} \left(\sum_{i=1}^{N} q_i \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{v}(k,\boldsymbol{l}) \cdot \boldsymbol{x}_i} \right) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{v}(k,\boldsymbol{l}) \cdot \boldsymbol{x}_j},
\end{aligned}$$

where we use the truncated Fourier series (26) and (27) and define $\boldsymbol{M} := (M_1, \tilde{\boldsymbol{M}}) \in 2\mathbb{N}^3$ as well as the vectors $\boldsymbol{v}(k, \boldsymbol{l}) := (k/L, \boldsymbol{l}/h) \in L^{-1}\mathbb{Z} \times h^{-1}\mathbb{Z}^2$.

The expressions in the inner brackets

$$\hat{S}(k, \boldsymbol{l}) := \sum_{i=1}^{N} q_i \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{v}(k,\boldsymbol{l}) \cdot \boldsymbol{x}_i}, \quad (k, \boldsymbol{l}) \in \mathcal{I}_{\boldsymbol{M}},$$

can be computed by an adjoint NFFT. This will be followed by $|\mathcal{I}_{\boldsymbol{M}}|$ multiplications with $\hat{b}_{|k|,\boldsymbol{l}}$ and completed by an NFFT to compute the outer summation over the indexes $(k, \boldsymbol{l}) \in \mathcal{I}_{\boldsymbol{M}}$. The proposed evaluation of $\phi^{\mathrm{p1,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p1},0}(\boldsymbol{x}_j)$ at the points $\boldsymbol{x}_j$, $j = 1, \ldots, N$, requires $\mathcal{O}(N + |\mathcal{I}_{\boldsymbol{M}}| \log |\mathcal{I}_{\boldsymbol{M}}|)$ arithmetic operations.

Again, using a matrix-vector notation we can write

$$\left(\phi^{\mathrm{p1,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p1},0}(\boldsymbol{x}_j)\right)_{j=1}^{N} \approx \tilde{\boldsymbol{A}} \boldsymbol{D} \tilde{\boldsymbol{A}}^{\mathrm{H}} \boldsymbol{q}, \tag{28}$$

where $\tilde{\boldsymbol{A}}$ denotes the matrix representation of the NFFT in three dimensions for the nodes $(x_{j,1}/L, \tilde{\boldsymbol{x}}_j/h) \in \mathbb{T}^3$, $\boldsymbol{D}$ is a diagonal matrix with entries $\hat{b}_{|k|,\boldsymbol{l}}$, $(k, \boldsymbol{l}) \in \mathcal{I}_{\boldsymbol{M}}$, and $\boldsymbol{q} = (q_1, \ldots, q_N)^{\top} \in \mathbb{R}^N$.

Relations to existing work

The Ewald formulas for 1d-periodic geometries were already proposed in[46]. Some Fourier based algorithms, like MMM1D[7], where already proposed. A method based on approximation Variant II from Section 2.1 is proposed in[36]. However, as mentioned in Section 2.1 this method suffers from the rather slow convergence rate in Fourier space. See also[53,10] for algorithms with higher complexity.

6 Fast Ewald Summation for 0d-Periodic (Open) Boundary Conditions

We consider a (not necessarily electrical neutral) system of N charges $q_j \in \mathbb{R}$ at positions $\boldsymbol{x}_j \in \mathbb{R}^3$, $j = 1, \ldots, N$. Under open boundary conditions the potential of each single particle j is defined by

$$\phi^{\mathrm{p0}}(\boldsymbol{x}_j) := \phi_{\{0\}^3}(\boldsymbol{x}_j) = {\sum_{i=1}^{N}}' \frac{q_i}{\|\boldsymbol{x}_{ij}\|} ,$$

i.e., we set $\mathcal{S} := \{0\}^3$ within the definition (3). This can be rewritten as

$$\phi^{\mathrm{p0}}(\boldsymbol{x}_j) = \phi^{\mathrm{p0,S}}(\boldsymbol{x}_j) + \phi^{\mathrm{p0,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p0,self}}(\boldsymbol{x}_j),$$

where for the splitting parameter $\alpha > 0$ we define the short range part

$$\phi^{\mathrm{p0,S}}(\boldsymbol{x}_j) := {\sum_{i=1}^{N}}' q_i \frac{\operatorname{erfc}(\alpha\|\boldsymbol{x}_{ij}\|)}{\|\boldsymbol{x}_{ij}\|},$$

the long range part

$$\phi^{\mathrm{p0,L}}(\boldsymbol{x}_j) := \sum_{i=1}^{N} q_i \Theta^{\mathrm{p0}}(\|\boldsymbol{x}_{ij}\|), \tag{29}$$

the self potential

$$\phi^{\mathrm{p0,self}}(\boldsymbol{x}_j) := -\frac{2\alpha}{\sqrt{\pi}} q_j ,$$

and the function $\Theta^{\mathrm{p0}}(r)$ is defined by

$$\Theta^{\mathrm{p0}}(r) = \begin{cases} \frac{2\alpha}{\sqrt{\pi}} & : r = 0, \\ \frac{\operatorname{erf}(\alpha r)}{r} & : \text{else.} \end{cases}$$

Using Variant III of Section 2, i.e., the fast summation approach, the function is approximated by a Fourier series. Similar as in the previous section we derive the fast algorithm now based on (29). The evaluation of the short range part $\phi^{\mathrm{p0,S}}(\boldsymbol{x}_j)$ is done by a direct evaluation again.

In the following we assume that we have $\|\boldsymbol{x}_{ij}\| \leq L =: h(1/2 - \varepsilon)$. As an example, if $\boldsymbol{x}_j \in L_1\mathbb{T} \times L_2\mathbb{T} \times L_3\mathbb{T}$, we can set $L := \sqrt{L_1^2 + L_2^2 + L_3^2}$. In order to approximate the long range part $\phi^{\mathrm{p0,L}}(\boldsymbol{x}_j)$ efficiently we consider the regularizations

$$K_{\mathrm{R}}(r) := \begin{cases} \Theta^{\mathrm{p0}}(r) & : |h^{-1}r| \leq 1/2 - \varepsilon, \\ K_{\mathrm{B}}(r) & : |h^{-1}r| \in (1/2 - \varepsilon, 1/2], , \\ K_{\mathrm{B}}(h/2) & : |h^{-1}r| > 1/2, \end{cases}$$

where the function $K_{\mathrm{B}}(\cdot) : [h/2 - h\varepsilon, h/2] \to \mathbb{R}$ is constructed such that $K_{\mathrm{R}}(\|\cdot\|) : h\mathbb{T}^3 \to \mathbb{R}$ is in the Sobolev space $\mathrm{C}^{p-1}(h\mathbb{T}^3)$, i.e., $K_{\mathrm{B}}(k, \cdot)$ fulfills the interpolation conditions

$$\frac{\mathrm{d}^j}{\mathrm{d}r^j} K_{\mathrm{B}}(h/2 - h\varepsilon) = \frac{\mathrm{d}^j}{\mathrm{d}r^j} \Theta^{\mathrm{p0}}(h/2 - h\varepsilon) \quad \text{for } j = 0, \ldots, p-1$$

as well as

$$\frac{\mathrm{d}^j}{\mathrm{d}r^j} K_{\mathrm{B}}(h/2) = 0 \quad \text{for } j = 1, \ldots, p-1. \tag{30}$$

Note that $K_{\mathrm{R}}(\|\cdot\|)$ is constant for all the points $\{\boldsymbol{y} \in h\mathbb{T}^3 : \|\boldsymbol{y}\| \geq h/2\}$. Therefore, the conditions (30) ensure smoothness of $K_{\mathrm{R}}(\|\cdot\|)$ in the points $\{\boldsymbol{y} \in h\mathbb{T}^3 : \|\boldsymbol{y}\| = h/2\}$. Furthermore, (30) does not include any restriction on the function value of $K_{\mathrm{R}}(h/2)$, since it does not influence the smoothness of $K_{\mathrm{R}}(\|\cdot\|)$. In Appendix C of[40] we show that an adopted version of Theorem 2.1 can be used to construct the regularizing functions $K_{\mathrm{B}}(k, \|.\|)$ as interpolation polynomials of degree $2p-2$. By our construction the function $K_{\mathrm{R}}(\|\cdot\|)$ is h-periodic in each direction and smooth, i.e., $K_{\mathrm{R}}(\|\cdot\|) \in \mathrm{C}^{p-1}(h\mathbb{T}^3)$.

To this end, we approximate the function

$$\Theta^{\mathrm{p0}}(\|\boldsymbol{y}\|) \approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{b}_{\boldsymbol{l}} \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l} \cdot \boldsymbol{y}/h}$$

for $\|\boldsymbol{y}\| \leq h/2 - h\varepsilon$ by a trigonometric polynomial. Thereby, we choose $\boldsymbol{M} = (M_1, M_2, M_3) \in 2\mathbb{N}^2$ large enough and compute the Fourier coefficients $\hat{b}_{\boldsymbol{l}}$ by

$$\hat{b}_{\boldsymbol{l}} := \frac{1}{|\mathcal{I}_{\boldsymbol{M}}|} \sum_{\boldsymbol{j} \in \mathcal{I}_{\boldsymbol{M}}} K_{\mathrm{R}}\left(\|\boldsymbol{j} \odot \boldsymbol{M}^{-1}\| h\right) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{j} \cdot (\boldsymbol{l} \odot \boldsymbol{M}^{-1})}.$$

In summary we obtain the following approximation for the long range parts

$$\begin{aligned} \phi^{\mathrm{p0,L}}(\boldsymbol{x}_j) &\approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{b}_{\boldsymbol{l}} \sum_{i=1}^{N} q_i \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l} \cdot \boldsymbol{x}_{ij}/h} \\ &= \sum_{\boldsymbol{l} \in \mathcal{I}_{\mathcal{M}}} \hat{b}_{\boldsymbol{l}} \left(\sum_{i=1}^{N} q_i \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l} \cdot \boldsymbol{x}_i/h} \right) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{l} \cdot \boldsymbol{x}_j/h}, \end{aligned}$$

where we use the truncated Fourier series. The expressions in the inner brackets

$$\hat{S}(\boldsymbol{l}) := \sum_{i=1}^{N} q_i \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l} \cdot \boldsymbol{x}_i/h}, \quad \boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}},$$

can be computed by an adjoint NFFT. This will be followed by $|\mathcal{I}_{\boldsymbol{M}}|$ multiplications with $\hat{b}_{\boldsymbol{l}}$ and completed by an NFFT to compute the outer summation with the complex

exponentials. The proposed evaluation of $\phi^{\mathrm{p1,L}}(\boldsymbol{x}_j) + \phi^{\mathrm{p1,0}}(\boldsymbol{x}_j)$ at the points $\boldsymbol{x}_j$, $j = 1, \ldots, N$, requires $\mathcal{O}(N + |\mathcal{I}_{\boldsymbol{M}}| \log |\mathcal{I}_{\boldsymbol{M}}|)$ arithmetic operations.

As for the different types of periodic boundary conditions, we also obtain a matrix-vector notation for the 0d-periodic case, which reads as

$$\left(\phi^{\mathrm{p0,L}}(\boldsymbol{x}_j)\right)_{j=1}^{N} \approx \tilde{\boldsymbol{A}}\boldsymbol{D}\tilde{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{q}, \tag{31}$$

where $\tilde{\boldsymbol{A}}$ denotes the matrix representation of the NFFT in three dimensions for the nodes $\boldsymbol{x}_j/h \in \mathbb{T}^3$, $\boldsymbol{D}$ is a diagonal matrix with entries $\hat{b}_{\boldsymbol{l}}$, $\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}$, and $\boldsymbol{q} = (q_1, \ldots, q_N)^{\top} \in \mathbb{R}^N$.

Relations to existing work

This method can be interpreted as nonequispaced convolution. For equispaced nodes the discrete convolution and its fast computation is typically realized by FFT exploiting the basic property $\mathrm{e}^{2\pi\mathrm{i}(y-x)} = \mathrm{e}^{2\pi\mathrm{i}y}\mathrm{e}^{-2\pi\mathrm{i}x}$. Following these lines, the method can be interpreted as "convolution at nonequispaced nodes" by Fourier methods as well, more precisely by the NFFT. This new method includes convolutions, e.g., with kernels of the form $1/\|x\|$. We remark that some FFT-accelerated Ewald[26,22] methods contain similar steps as the fast summation based on NFFT. A method based on approximation Variant II from Section 2.1 is proposed in[37]. However, as mentioned in Section 2.1 this method suffers from the rather slow convergence rate in Fourier space.

7 Conclusion

With this tutorial we provide an overview of the NFFT based fast Ewald summation for all kinds of mixed periodic boundary conditions. The main advantage of our approach is that all presented algorithms have a common structure. More precisely, the short range parts of the potentials are always computed directly and the long range parts are computed by an adjoint NFFT, a point-wise multiplication in Fourier space and, again, an NFFT in three dimensions, see the matrix vector notation for the 3d-periodic case in (12), for the 2d-periodic case in (20), for the 1d-periodic case in (28), and for the 0d-periodic case in (31). Non-periodic boundary conditions are handled via a combination of the corresponding mixed-periodic Ewald formulas and the concept of NFFT based fast summation. Thereby, we embed the non-periodic functions of the mixed-periodic Ewald formulas into smooth periodic functions and obtain rapidly convergent Fourier approximations. Note that this approach also includes the 0d-periodic (open) case.

Since all algorithms depend on common building blocks, code and algorithm improvements can be realized individually on the more elementary submodules. For example, a MPI-based parallel NFFT algorithm was proposed in[45] and is publicly available[42]. The parallelization of the NFFT module led to a parallel version of all the above mentioned Ewald summation methods at once.

Following the naming scheme of the particle-particle particle-mesh ($\mathrm{P^3M}$) method, our proposed framework is called particle-particle NFFT ($\mathrm{P^2NFFT}$), since the short range particle-particle interactions are computed in the same way as in $\mathrm{P^3M}$ algorithms, while the long range particle-mesh part is computed by NFFTs.

In order to rate the very good performance of the proposed algorithms, we compared the method to the P^{2}NFFT method for 3d-periodic systems[45] as well as to the method proposed in[34] by considering similar numerical examples, see[40]. Note that the P^{2}NFFT algorithm is highly optimized, publicly available[1], and recently compared to other methods, such as the P^{3}M method, the fast multipole method and multigrid based methods, see[5] and the references therein.

Acknowledgments

The authors gratefully acknowledge support by the German Research Foundation (DFG) project PO 711/12-1.

References

1. ScaFaCoS - Scalable Fast Coloumb Solvers. `http://www.scafacos.de`.
2. M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions*. National Bureau of Standards, Washington, DC, USA, 1972.
3. R. P. Agarwal and P. J. Y. Wong. *Error inequalities in polynomial interpolation and their applications*, volume 262 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1993.
4. A. Arnold, J. de Joannis, and C. Holm. Electrostatics in periodic slab geometries. I. *J. Chem. Phys.*, 117:2496, 2002.
5. A. Arnold, F. Fahrenberger, C. Holm, O. Lenz, M. Bolten, H. Dachsel, R. Halver, I. Kabadshow, F. Gähler, F. Heber, J. Iseringhausen, M. Hofmann, M. Pippig, D. Potts, and G. Sutmann. Comparison of scalable fast methods for long-range interactions. *Phys. Rev. E*, 88:063308, 2013.
6. A. Arnold and C. Holm. MMM2D: A fast and accurate summation method for electrostatic interactions in 2D slab geometries. *Comput. Phys. Commun.*, 148:327 – 348, 2002.
7. A. Arnold and C. Holm. MMM1D: A method for calculating electrostatic interactions in one-dimensional periodic geometries. *J. Chem. Phys.*, 123:144103, 2005.
8. G. Beylkin. On the fast Fourier transform of functions with singularities. *Appl. Comput. Harmon. Anal.*, 2:363 – 381, 1995.
9. A. Bródka. Ewald summation method with electrostatic layer correction for interactions of point dipoles in slab geometry. *Chem. Phys. Lett.*, 400:62 – 67, 2004.
10. A. Bródka and P. Sliwinski. Three-dimensional Ewald method with correction term for a system periodic in one direction. *J. Chem. Phys.*, 120:5518 – 5523, 2004.
11. T. Darden, D. York, and L. Pedersen. Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.*, 98:10089–10092, 1993.
12. S. W. de Leeuw, J. W. Perram, and E. R. Smith. Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constants. *Proc. Roy. Soc. London Ser. A*, 373:27 – 56, 1980.
13. M. Deserno and C. Holm. How to mesh up Ewald sums. I. A theoretical and numerical comparison of various particle mesh routines. *J. Chem. Phys.*, 109:7678 – 7693, 1998.

14. A. Dutt and V. Rokhlin. Fast Fourier transforms for nonequispaced data. *SIAM J. Sci. Stat. Comput.*, 14:1368 – 1393, 1993.
15. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. A smooth particle mesh Ewald method. *J. Chem. Phys.*, 103:8577 – 8593, 1995.
16. P. P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.*, 369:253–287, 1921.
17. M. Fenn and G. Steidl. Fast NFFT based summation of radial functions. *Sampl. Theory Signal Image Process.*, 3:1 – 28, 2004.
18. J. A. Fessler and B. P. Sutton. Nonuniform fast Fourier transforms using min-max interpolation. *IEEE Trans. Signal Process.*, 51:560 – 574, 2003.
19. D. Frenkel and B. Smit. *Understanding molecular simulation: From algorithms to applications*. Academic Press, 2002.
20. M. Frigo and S. G. Johnson. FFTW, C subroutine library. `http://www.fftw.org`, 2009.
21. L. Greengard and J.-Y. Lee. Accelerating the nonuniform fast Fourier transform. *SIAM Rev.*, 46:443 – 454, 2004.
22. M. Griebel, S. Knapek, and G. Zumbusch. *Numerical simulation in molecular dynamics*, volume 5 of *Texts in Computational Science and Engineering*. Springer, Berlin, 2007.
23. A. Grzybowski, E. Gwóźdź, and A. Bródka. Ewald summation of electrostatic interactions in molecular dynamics of a three-dimensional system with periodicity in two directions. *Phys. Rev. B*, 61:6706–6712, 2000.
24. F. E. Harris. Incomplete Bessel, generalized incomplete gamma, or leaky aquifer functions. *J. Comput. Appl. Math.*, 215:260 – 269, 2008.
25. F. Hedman and A. Laaksonen. Ewald summation based on nonuniform fast Fourier transform. *Chem. Phys. Lett.*, 425:142 – 147, 2006.
26. R. W. Hockney and J. W. Eastwood. *Computer simulation using particles*. Taylor & Francis, Inc., Bristol, PA, USA, 1988.
27. P. H. Hünenberger. Lattice-sum methods for computing electrostatic interactions in molecular simulations. In *Simulation and theory of electrostatic interactions in solution*, volume 17, pages 17 – 83. ASCE, 1999.
28. D. Huybrechs. On the Fourier extension of nonperiodic functions. *SIAM J. Numer. Anal.*, 47(6):4326 – 4355, 2010.
29. J. Keiner, S. Kunis, and D. Potts. NFFT 3.0, C subroutine library. `http://www.tu-chemnitz.de/~potts/nfft`.
30. J. Keiner, S. Kunis, and D. Potts. Using NFFT3 - a software library for various nonequispaced fast Fourier transforms. *ACM Trans. Math. Software*, 36:Article 19, 1 – 30, 2009.
31. S. Kunis and S. Kunis. The nonequispaced FFT on graphics processing units. *PAMM, Proc. Appl. Math. Mech.*, 12, 2012.
32. S. Kunis and D. Potts. Time and memory requirements of the nonequispaced FFT. *Sampl. Theory Signal Image Process.*, 7:77 – 100, 2008.
33. D. Lindbo and A.-K. Tornberg. Spectral accuracy in fast Ewald-based methods for particle simulations. *J. Comput. Phys.*, 230:8744 – 8761, 2011.
34. D. Lindbo and A.-K. Tornberg. Fast and spectrally accurate Ewald summation for 2-periodic electrostatic systems. *J. Chem. Phys.*, 136:164111, 2012.

35. M. Lyon. Approximation error in regularized SVD-based Fourier continuations. *Appl. Numer. Math.*, 62(12):1790 – 1803, 2012.
36. P. Minary, J. A. Morrone, D. A. Yarne, M. E. Tuckerman, and G. J. Martyna. Long range interactions on wires: a reciprocal space based formalism. *J. Chem. Phys.*, 121:11949 – 11956, 2004.
37. P. Minary and M. E. Tuckerman. A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *J. Chem. Phys.*, 110:2810 – 2821, 1999.
38. P. Minary, M. E. Tuckerman, K. A. Pihakari, and G. J. Martyna. A new reciprocal space based treatment of long range interactions on surfaces. *J. Chem. Phys.*, 116:5351 – 5362, 2002.
39. F. Nestler. Automated parameter tuning based on RMS errors for nonequispaced FFTs. *Preprint 2015-01, Faculty of Mathematics, Technische Universität Chemnitz*, 2015.
40. F. Nestler, M. Pippig, and D. Potts. Fast Ewald summation based on NFFT with mixed periodicity. *J. Comput. Phys.*, 2015.
41. M. Pippig. PFFT, Parallel FFT subroutine library. `http://www.tu-chemnitz.de/~mpip/software.php`, 2011.
42. M. Pippig. PNFFT, Parallel Nonequispaced FFT subroutine library. `http://www.tu-chemnitz.de/~mpip/software.php`, 2011.
43. M. Pippig. PFFT - An extension of FFTW to massively parallel architectures. *SIAM J. Sci. Comput.*, 35:C213 – C236, 2013.
44. M. Pippig and D. Potts. Particle simulation based on nonequispaced fast Fourier transforms. In G. Sutmann, P. Gibbon, and T. Lippert, editors, *Fast Methods for Long-Range Interactions in Complex Systems*, IAS-Series, pages 131 – 158, Jülich, 2011. Forschungszentrum Jülich.
45. M. Pippig and D. Potts. Parallel three-dimensional nonequispaced fast Fourier transforms and their application to particle simulation. *SIAM J. Sci. Comput.*, 35:C411 – C437, 2013.
46. M. Porto. Ewald summation of electrostatic interactions of systems with finite extent in two of three dimensions. *J. Phys. A*, 33:6211 – 6218, 2000.
47. D. Potts and G. Steidl. Fast summation at nonequispaced knots by NFFTs. *SIAM J. Sci. Comput.*, 24:2013 – 2037, 2003.
48. D. Potts, G. Steidl, and M. Tasche. Fast Fourier transforms for nonequispaced data: A tutorial. In J. J. Benedetto and P. J. S. G. Ferreira, editors, *Modern Sampling Theory: Mathematics and Applications*, pages 247 – 270, Boston, MA, USA, 2001. Birkhäuser.
49. G. Steidl. A note on fast Fourier transforms for nonequispaced grids. *Adv. Comput. Math.*, 9:337 – 353, 1998.
50. T. Volkmer. OpenMP parallelization in the NFFT software library. *Preprint TU Chemnitz*, Preprint 7, 2012.
51. Y.-L. Wang, F. Hedman, M. Porcu, F. Mocci, and A. Laaksonen. Non-uniform FFT and its applications in particle simulations. *App. Math.*, 5:520 – 541, 2014.
52. A. H. Widmann and D. B. Adolf. A comparison of Ewald summation techniques for planar surfaces. *Comput. Phys. Commun.*, 107:167 – 186, 1997.
53. I.-C. Yeh and M. L. Berkowitz. Ewald summation for systems with slab geometry. *J. Chem. Phys.*, 111(7):3155 – 3162, 1999.

Modern Trends in Hardware Development

Dirk Pleiter

Jülich Supercomputing Centre,
Forschungszentrum Jülich, 52425 Jülich, Germany

E-mail: d.pleiter@fz-juelich.de

The need for high-performance computers (HPC) which are capable of solving large-scale numerical problems is growing. In this lecture we provide an overview on current trends in development of HPC architectures and hardware technologies. These will in future allow for further significant enhancement of performance. However, there are technological trends which have significant impact on the architectural design space. These should be taken into account by application developers in order to be prepared for efficient use of future HPC architectures.

1 Introduction

Different studies indicate a growing need for supercomputing resources to address challenging problems in science and engineering, at some point reaching the need for exascale facilities.[1,2] One of the key areas which was identified is material science, chemistry and nano-science. Increase of scalable compute resources are necessary to enable, e.g., complete simulations of entire systems. Increasing the resources, which are today available for single pioneering calculations on high-end capability systems, by 2-3 orders of magnitude would allow to perform many such simulations concurrently, i.e. high-throughput runs for several systems and many different parameters become possible on large capacity systems.

Since at least 2 decades, research and development on HPC architectures and technologies resulted in an exponential growth in available computing resources. Since 1993 the performance of high-end HPC systems can be compared in terms of throughput of double-precision floating-point operations while executing the High-Performance LINPACK (HPL) benchmark, i.e. while solving a dense linear set of equations. The results are used for ranking the systems in the Top500 List.[3] In Fig. 1 we show the performance achieved as a function of time for the number one systems as well as the aggregate performance of the top 10 and 100 systems. For the latter case we also show how the performance since 1993 would have evolved assuming a doubling of the performance every 14 months. In the context of this observed scaling, often the term "Moore's Law" is used, which is misleading in several respects. The term was initially introduced to describe the observed increase in density of transistors within integrated circuits.[4] As we will see later this does have a strong impact on improving performance, but there are several other architectural, technological as well as commercial factors that drive the growth in overall system performance. Therefore, this behaviour should be considered as a mere phenomenological observation rather then the consequence of more fundamental mechanisms.

In this lecture we will have a deeper look into this trend of performance increase and discuss some of the underlying technologies. Before we explore key technology trends in section 3, we will provide a model based framework for analysing computer architectures and performance in section 2. On this basis we will look into processor, accelerator and

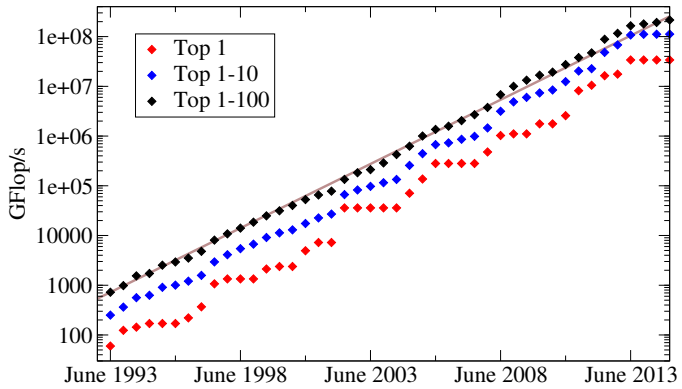


Figure 1: Evolution of the aggregate performance achieved for the top 1, 10 and 100 systems from the Top500 List[3] over time. The line shows how the aggregate performance of the top 100 systems would have developed from June 1993 on assuming a doubling of the performance every 14 months.

network technologies in the sections 4, 5 and 6, respectively. To connect back to applications for the research fields, which are topic of this Jülich School in Computational Science, section 7 provides a digression where we analyse requirements for a specific computational fluid-dynamics method, namely the Lattice-Boltzmann method. This will be followed by a presentation of two examples for state-of-the-art supercomputer architectures in section 8.

2 Abstract Machines and Performance Models

State-of-the-art HPC architectures and many of their components, e.g. computing devices like processors or network devices, have become extremely complex. To keep the level of complexity for application developers (but also HPC system architects) manageable it is mandatory to introduce suitable levels of abstractions. *Abstract machine models* are becoming a commonly used vehicle for enabling so-called co-design processes, where system architects and technology engineers work together with developers of applications and algorithms for designing new system architectures as well as algorithms and implementations of applications that can efficiently exploit these new architectures. Recently, Ang et al.[5] suggested the use of what they call *proxy architecture models*, which are abstract models of possible system architectures that are parametric. The resulting parameter space can then be explored using, e.g., performance models to optimize the architecture for a given set of applications. Architectural parameters include, e.g., throughput of floating-point operations or memory bandwidth and capacity.

Performance models have a long history in computer science, but are also becoming more widely used in computational science. Ideally, the effort for creating application specific performance models for a given system architecture is significantly below that of creating an optimized implementation. Models thus allow for a quick assessment of the performance one could hope to achieve for a given architecture and algorithm. Hoefler et al.[6] suggested application developers to use models for all phases of software design following the waterfall model. The later refers to a design process which sequentially progresses through multiple phases. During an initial analysis phase performance models

can help to select the right algorithms, while during the later implementation phase such models can become the basis for making trade-off decisions. Even during the final maintenance phase performance models can be useful, e.g. to identify performance problems for particular input parameter choices.

Performance models are always based on an underlying machine model. The performance predictions typically depend on the parameters of such (often implicitly assumed) models.[a] Predictions obviously also depend on the application or algorithm specific parameters, in particular problem size related parameters. Suitable choices of such application specific parameters are often difficult to fix, because this requires a deeper analysis of the research goals. But they nevertheless can have huge impact on the achievable performance.[b]

The choice of an abstract machine model considered in this lecture is based on the assumption that data transport is a major performance limiting factor for today's computer architectures. In this model we consider a computer architecture being composed of two types of devices:

- **Storage devices**: A storage device allows to hold a certain amount of data and is characterized by its storage capacity C. An example for a storage device is the processor's main memory that features a given capacity C_{mem}.

- **Transport or processing devices**: A transport device connects storage devices to facilitate data transport between these devices. Processing devices are very similar but additionally provide data processing capabilities. The key parameter for such devices is bandwidth (or throughput) B. Note that a processing device might connect a storage device with itself. One example for such a processing device is an arithmetic unit capable of performing floating-point operations at a rate B_{fp} in units of Flop/s.

This simple model can easily be mapped onto a graph where the nodes and edges represent storage and transport or processing devices, respectively. In Fig. 2 (left) we show a model for a very simple processor architecture comprising two storage devices, an external main memory (MEM) and a register file (RF),[c] and two transport/processing devices: a memory bus and an arithmetic unit. During program execution data is moved via the memory bus, a transport device connecting external main memory and register file, into the processor. Once the data is available in the register file the arithmetic pipeline can take this as input and write the results of the arithmetic operation back to the register file, i.e. the same storage device.

In Fig. 2 (right) we provide an example for adapting the level of abstraction. It shows a model for an architecture comprising multiple processors interconnected in a ring topology by connecting pairs of processors through a network link. In this model all details of the node's processor architecture are omitted.

---

[a] In some cases the underlying machine model is not parametric. E.g., semi-empirical performance models determine machine dependent parameters from fits to application performance measurements.

[b] The Top500 List is often criticised for not fixing the size of the linear system to be solved. Any supplier of HPC solutions is free to choose a problem size for which performance for a given system is maximized.

[c] The term *register file* refers to a set of registers. A register file features one (or more) input (output) ports for writing data to (reading data from) any of the registers. Register files exist in all processor architectures currently used for HPC systems.

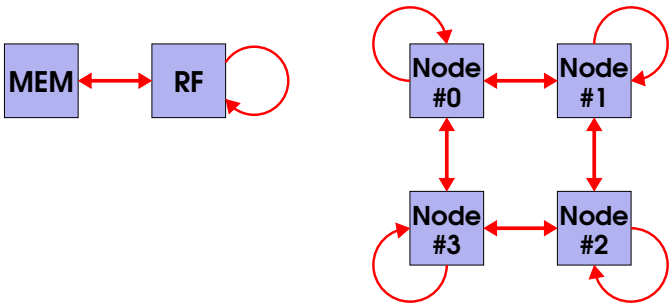


Figure 2: Abstract machine model for a single processor (left) and a multi-node (right) system. More details are explained in the text.

Based on this abstract machine model we introduce a simple performance model that is based on the *information exchange* function $I^k_{x,y}(W)$.[7] For a given computational task identified by the index k this function gives the amount of information which needs to be exchanged between the storage devices x and y. The amount of information depends on the problem size W. Let us consider an example from the Basic Linear Algebra Subroutines (BLAS) library which is widely used, namely the SAXPY function. This function implements the computational task

$$\vec{y} \leftarrow \alpha\,\vec{x} + \vec{y}, \tag{1}$$

where α is a scalar and $\vec{x}$, $\vec{y}$ are arrays of single-precision floating-point numbers. The problem size is parametrized by the array length which we denote with N. An implementation of this function has to load all elements of the input arrays $\vec{x}$, $\vec{y}$ from memory to register file and store the output array $\vec{y}$. For each array element the arithmetic unit has to perform one multiplication and one addition. We thus obtain the following information exchange functions:

$$I_{\text{mem}} = (2+1)\cdot N \cdot 4 \text{ Byte}, \tag{2}$$

$$I_{\text{fp}} = 2 \cdot N \text{ Flop}. \tag{3}$$

We assume that the scalar variable can be kept in the processor while executing this task. For sufficiently large N we can ignore loading α.

To estimate the performance we use the *latency-bandwidth model*. In this model it is assumed that the time required to transport or process a certain amount of data I can be parametrized linearly using the following ansatz:

$$t = \lambda + \frac{I}{\beta}, \tag{4}$$

where λ refers to the start-up latency, i.e. the time from start of an operation until the first data arrives. β denotes the asymptotic bandwidth, i.e. the bandwidth which is observed for large problem sizes where the start-up latency becomes negligible. This asymptotic bandwidth β must be smaller than the nominal hardware bandwidth B, i.e. $\beta \leq B$.

In case of the SAXPY function the time needed to load and store data as well as to process the data can be estimated as follows:

$$t_{\mathrm{mem}} = \lambda_{\mathrm{mem}} + \frac{(2+1)\cdot N \cdot 4\ \mathrm{Byte}}{\beta_{\mathrm{mem}}} \tag{5}$$

$$t_{\mathrm{fp}} = \lambda_{\mathrm{fp}} + \frac{2\cdot N\ \mathrm{Flop}}{\beta_{\mathrm{fp}}}. \tag{6}$$

If we make the assumption that load or store and arithmetic operations can be executed concurrently, then the model predicts the total execution time to be given by

$$t = \max(t_{\mathrm{mem}}, t_{\mathrm{fp}}). \tag{7}$$

For $t_{\mathrm{mem}} < t_{\mathrm{fp}}$ the performance is considered to be *compute bound* or otherwise *memory bound*.

The ratio of arithmetic operations versus the amount of data, that needs to be moved from the relevant storage device, e.g. external main memory, to the arithmetic units, is called *Arithmetic Intensity*.[8] In our framework we can define the Arithmetic Intensity as follows:

$$\mathrm{AI} = \frac{I_{\mathrm{fp}}}{I_{\mathrm{mem}}}. \tag{8}$$

For the SAPY function we thus have $\mathrm{AI} = 1\,\mathrm{Flop}/6\,\mathrm{Byte}$.

Let us assume the start-up latencies to be negligible.[d] We then can use the Arithmetic Intensity to express the predicted sustained floating-point performance b_{fp} of the SAXPY function as follows:

$$b_{\mathrm{fp}} = \frac{I_{\mathrm{fp}}}{t} = \min\left(\frac{I_{\mathrm{fp}}}{t_{\mathrm{fp}}}, \frac{I_{\mathrm{fp}}}{t_{\mathrm{mem}}}\right) \simeq \min\left(\beta_{\mathrm{fp}}, \mathrm{AI}\cdot\beta_{\mathrm{mem}}\right) \leq \min\left(B_{\mathrm{fp}}, \mathrm{AI}\cdot B_{\mathrm{mem}}\right). \tag{9}$$

Through this simple arithmetic transformation we obtained an upper limit (roof) for the achievable performance for a given computational device, which is characterized by the hardware parameters B_{fp} and B_{mem}, as a function of the Arithmetic Intensity, which is a feature of the given computational task. This approach to assess the maximum attainable performance is called *Roofline Model*.[9]

For small AI the computational task is memory bound. With increasing AI the attainable performance increases until $\mathrm{AI} = B_{\mathrm{mem}}/B_{\mathrm{fp}}$, when the point of maximum performance is reached. For larger AI the task becomes compute bound for the given compute device. In Fig. 3 we show the rooflines for different processor and accelerator devices. From this plot one can easily see that the performance of the SAXPY function with $\mathrm{AI} \simeq 0.2\,\mathrm{Flop/Byte}$ is bound by the memory bandwidth for all compute devices which we will discuss in more detail in the next sections.

3 Technology Trends

For about two decades the CMOS technology, which is the basis for all computational devices considered here, could be improved while keeping electrical fields roughly constant.

[d] This assumption will hold for sufficiently large array length N.

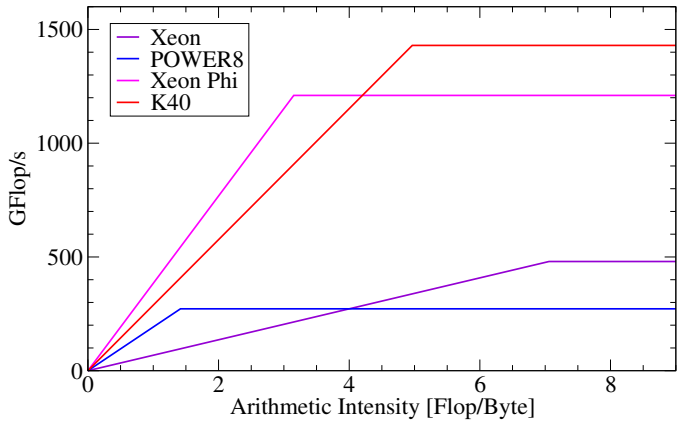


Figure 3: Results for the roofline model for the processors and accelerator devices discussed in in more detail in the sections 4 and 5. The hardware parameters are listed in Table 1 and 2.

This could be achieved by increasing the transistor density (following the above mentioned Moore's scaling[4]), increasing the switching speed and reducing the supply voltage. These CMOS technology improvements, following the so-called *Dennard scaling*,[10] allowed to improve performance in two ways:

1. The increase in transistor density enabled more complex designs comprising more or more complex execution units. This allowed to increase the number of concurrently executed operations.

2. Increasing switching frequency allowed for higher clock speeds and as a result higher throughput of instructions and thus again more operations per time unit.

The end of Dennard scaling was reached about 10 years ago when supply voltage reached about 1 V and further voltage reduction became a challenge.[11] As a consequence it became more beneficial to only increase transistor density, i.e. Moore's scaling still holds, at constant clock speed. Dennard and collaborators later published an analysis where they concluded that a reduction down to about 0.5 V would be beneficial in terms of power efficiency.[12] But they did foresee a significant reduction of the clock speed.

This technology trend has significant impact on the scaling of applications. With scaling we mean that if machine performance, e.g. the floating-point performance B_{fp}, increases by a factor S then a similar speed-up is observed for the sustained performance b_{fp}. If an increase of B_{fp} is due to an increase of the number of concurrently executed floating-point operations then the application likely also has to become more parallel in order to allow for a sufficient number of operations to be executed independently.

A more fundamental rule affecting computer architectures at different levels is *Rent's rule*, which relates the number of logic elements (or gates) G and the number of edge connections (or terminals) T:[13]

$$T = k\, G^p, \tag{10}$$

where the Rent coefficient k gives the average number of connections per logic element and the *Rent exponent* p determines the scaling. The number of logic elements G relates

to the computational performance, while the number of edge connections T is related to the data transport capabilities. As often $p \ll 1$ it follows that it becomes challenging to keep communication and computational performance balanced when the number of logic elements increases.[14] Efficient packaging therefore is an increasingly important aspect of modern architectures and with increasing system size communication becomes a limiting factor.[15] Minimizing data transport by exploiting data locality properties of applications, which is already crucial today, will become even more important.

Using the terminology of the abstract machine architecture introduced in the previous section, modern processor and system architectures comprise various types of storage devices. They differ significantly in terms of bandwidth B and capacity C and are typically arranged hierarchically, with fast but small (slow but large) storage devices being closer to (further away from) the processing pipelines. With faster computing devices, like the so-called accelerators (see section 5), the memory hierarchy becomes deeper due to the need for particularly fast memory.

This second important technology trend requires the application developers to further improve on data locality. This is even more challenging as current programming environments do provide only few ways of expressing information about data locality, i.e. the burden of exploiting data locality largely remains with the developer.

4 Processor Architectures

In this section we discuss two processors which are suitable for state-of-the-art HPC systems and which have become available in 2014, namely a new Intel Xeon processor of the Haswell generation and the POWER8 processor from IBM. Intel Xeon processors currently dominate in the market of high-end HPC systems as more than 80 % of the systems listed in the Top500[3] List (as of November 2014) are based on processors of this type. A selection of hardware parameters of the processors considered in this section are listed in Table 1.

Processors, like the ones presented here, feature significant amounts of parallelism at different levels:

- *Core parallelism*: The processors comprise 10 cores or more. Each core can execute independent sets of instructions, but share parts of the memory hierarchy like the interfaces towards the external memory.

- *Instruction Level Parallelism (ILP)*: Each core can at each clock cycle execute multiple instructions concurrently, e.g. multiple arithmetic instructions plus load or store instructions.

- *Data-parallel instructions*: Vector or SIMD[e] instructions allow to perform the same operation(s) on multiple operands. This allows, e.g., to add two short vectors of floating-point numbers using just a single instruction.

Combining all this different levels of parallelism, processors can perform several hundreds of double-precision floating-point operations per clock cycle.

[e] SIMD stands for Single Instruction, Multiple Data.

Application programmers often have to cope with an additional level of parallelism as processor architectures provide hardware support for executing multiple threads, i.e. the ability to process independent sets of instructions. This hardware feature is often called Simultaneous Multi-Threading (SMT). Each core of the POWER8 processor supports 8-way SMT, i.e. such a processor with 10 cores can execute 80 threads concurrently. While all threads running on one core share most of the hardware resources, using multiple threads can help improving the utilization of these resources. For instance, if execution of one thread is stopped while waiting for a load instruction to complete then another thread may be ready for executing arithmetic instructions.

While floating-point performance B_{fp} continues to increase at significant rates, efficient exploitation of this performance becomes more challenging as memory bandwidth B_{mem} increases at a much lower rate. In Fig. 3 the roofline curves for both processors are shown. Both differ significantly in terms of minimum Arithmetic Intensity that would be theoretically required for maximum floating-point performance being attainable. The slope of the left part of the roofline is given by the ratio $B_{\mathrm{mem}}/B_{\mathrm{fp}}$.

The roofline plotted in Fig. 3, however, assumes that all input to and all output from the processing pipelines needs to be moved over the memory bus. This is often not the case as in practice most applications feature a significant amount of data locality. This means that most of the data that was loaded into the processor will be accessed again shortly thereafter. If the data can be kept in a cache until it is re-used it does not have to be loaded over the memory interface again. Caches feature much higher bandwidth, in particular the caches which are closest to the processing pipelines, typically called L1 cache. The available memory capacity is, however, significantly smaller as can be seen from Table 1 by comparing C_{L1d} and C_{mem}.[f] In this table we also list the maximum bandwidth $B_{\mathrm{ld/st}}$ that in theory could be achieved for the maximum possible throughput of load and store instructions. As each core has its own private L1 cache, the aggregate bandwidth for reading from or writing to these caches scales with the number of cores.

	Intel E5-2680 v3	IBM POWER8
Core clock speed [GHz]	2.5	3.4
Number of cores	12	10
B_{fp} [Flop/cycle]	192	80
B_{fp} [GFlop/s]	480	272
$B_{\mathrm{ld/st}}$ [GByte/s]	$12 \cdot 240$	$10 \cdot 218$
C_{L1d} [kiByte]	$12 \cdot 64$	$10 \cdot 64$
B_{mem} [GByte/s]	68	192
C_{mem} [GiByte]	$O(100)$	$O(100 \ldots 1000)$

Table 1: Key hardware parameters of processors suitable for HPC architectures.

[f] We refer to an L1d cache, which is used only for data, to distinguish it from the L1i cache, which is used only for instructions.

5 Accelerators

By using simpler core architectures as well as lower clock frequencies it is possible to increase the parallelism of compute devices even more. A typical example for such architectures are graphics processing units (GPU), which are meanwhile extensively used for numerical calculations. Simulations of field theories were among the first applications that did exploit the performance of GPUs.[16] Recently, the Many Integrated Core (MIC) architecture has been introduced by Intel. The architecture is much more similar to traditional processors but with the number of (simple) cores being significantly increased and particularly wide SIMD pipelines integrated.

As of today, these devices are not operated stand-alone but in combination with processors. The main program is executed on the processor and only selected functions are off-loaded to these so-called accelerators. In Table 2 we compare the performance parameters of currently available accelerators which are suitable for HPC systems. By comparing these hardware parameters with those for the processors listed in Table 1 the following can be observed:

- The core clock frequency used for accelerators is 2-4 times smaller compared to processors.
- The number of double-precision operations per clock cycle is $O(1000)$, i.e. 6-17 times larger then for the considered processors.
- As a consequence, significantly more floating-operations can be executed per time unit.

In Fig. 3 we also show the rooflines for the accelerator devices discussed in this section. Again a high Arithmetic Intensity is required before reaching the point where performance is expected to be mainly limited by the maximum throughput of floating-point instructions. Compared to the performance of these devices the aggregate cache and other on-chip memory capacity is small. Therefore, exploiting data locality is much harder compared to the processors discussed in the previous section.

In Fig. 4 we show a typical single-socket node architecture with an accelerator device. Accelerators are attached via an I/O interface to the processor. This link is relatively slow compared to the bandwidth available on the memory interfaces of both, processor and accelerator. To avoid this becoming a performance bottleneck the amount of data transferred

	Intel Xeon Phi 7120	NVIDIA K40
Core clock speed [GHz]	1.24	0.75
Number of cores	61	–
B_{fp} [Flop/cycle]	976	1920
B_{fp} [GFlop/s]	1210	1430
B_{mem} [GByte/s]	384	288
C_{mem} [GiByte]	8-16	12

Table 2: Key hardware parameters of accelerator devices suitable for HPC architectures.

over this link must be minimized. Another strategy to hide the effects of this slow connection is to overlap transfer with other computations.

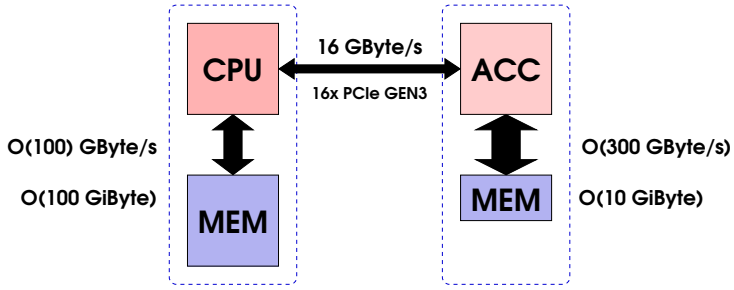


Figure 4: Typical single-socket node architecture with a processor (CPU) and an accelerator (ACC), like those shown in Table 1 and 2, respectively. Both, processor and accelerator have their own main memory (MEM) attached. The dashed line indicates the separated memory coherence domains of processor and accelerator.

6 High-Performance Networks

Using one or more accelerator devices it is possible to realize compute nodes that feature a floating-point performance of $O(1 \ldots 10)$ TFlop/s. To increase the performance even more requires a network that interconnects many compute nodes.

A network typically interconnects different types of nodes, namely compute nodes and switches. All high-performance networks are based on point-to-point connections between these nodes. Messages which are sent from one node thus usually have to traverse multiple links to reach the receiving node. The performance characteristics of such networks are largely determined by the link technology, network topology as well as the switch and routing technology.

The performance of technologies which are used for point-to-point communication can usually be well described by a latency-bandwidth model as shown in Eq. (4). The start-up latency refers to the time from start of the transfer until arrival of the first data at the destination. For low-latency network technologies $\lambda \simeq 1\,\mu$sec, i.e. $O(3000)$ processor clock cycles. While it becomes more difficult to reduce the start-up latency, link bandwidth continues to increase. Commonly used link technologies provide a bandwidth $B_{\text{net}} \simeq O(20 \ldots 50)$ Gbit/s which is about to increase to $B_{\text{net}} = 100$ Gbit/s.

Also the network topology has a major impact on the overall network performance. Depending on the communication patterns a network topology may be more or less suitable. There are different performance relevant metrics which characterise a network topology. In this lecture we limit us to the following metrics:

- *Network diameter*: The diameter is the maximum distance between 2 nodes. The larger the diameter the more hops may be required to perform a point-to-point communication between nodes and the start-up latency for communication will thus increase. This also affects the performance of collective operations where a small

amount of data has to be collected from all nodes (e.g., in case of a global sum) and/or distributed to all nodes (e.g., in case of a broadcast operation).

- *Bi-section bandwidth*: The bi-section bandwidth refers to the aggregate bandwidth of all links connecting two equal halves of the machine. This metric is important for all-to-all communication patters where large amounts of data are exchanged. In this case all nodes communicate with all other nodes and therefore all nodes of one half of the system have to communicate with all nodes of the other half of the system.

The following network topologies can be most frequently found in current high-end HPC architectures:

- *Fat-tree topology*: In a tree topology the non-leaf nodes are switches and leaf nodes are compute nodes. The advantage of tree topologies is that the diameter grows as $\log(N)$, where N is the number of compute nodes. To avoid small bi-section bandwidth the fat-tree topology was proposed.[17] In this topology the number of links towards the root level is increased. A disadvantage of fat-tree topologies is the large number of switches required when N becomes large, which results in high costs.

- *d-dimensional torus topology*: A torus topology is a popular choice for architectures with a very large number of compute nodes N. Network diameter and bi-section bandwidth scale as $N^{1/d}$ and N^{d-1}, respectively. While larger d improves these properties of the network, the costs also become large, e.g., because widely separated nodes need to be connected using long cables. In today's systems one typically finds $3 \leq d \leq 5$.

- *Dragonfly topology*:[18] This topology features very low diameter while compromising on bi-section bandwidth. The dragonfly topology is a multilevel topology where at the lowest level a set of compute nodes is connected to a high-radix switch. A group of such switches is then interconnected using an all-to-all topology. Yet another level can be formed by interconnecting these groups of switches. The network diameter scales as $N^{1/k}$, where k is the switch radix, i.e. the scaling properties are better compared to tree topologies.

7 Digression: Lattice-Boltzmann Method

The Lattice-Boltzmann Method (LBM) is an approach to computational fluid-dynamics which requires HPC resources to investigate problems of interest. We want to use LBM as a non-trivial example to review different topics related to modern HPC architectures and technologies presented in the previous sections.

The computations follow an iterative update scheme where for each location $\vec{x}_i$ the following expression has to be evaluated:[g]

$$f_\alpha(\vec{x}_i + \vec{e}_{\alpha,i}\,\Delta t, t + \Delta t) = f_\alpha(\vec{x}_i, t) - \frac{\Delta t}{\tau}\left[f_\alpha(\vec{x}_i, t) - f_\alpha^{(\mathrm{eq})}\left(\rho(\vec{x}_i, t), \vec{u}_i(\vec{x}_i, t)\right)\right]. \quad (11)$$

[g] We largely follow the notation of Feichtinger et al.[19]

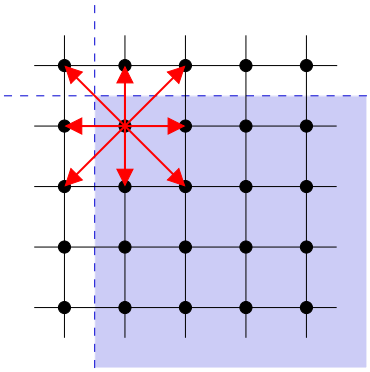


Figure 5: Graphical representation of the velocities in the Lattice-Boltzmann Method model D2Q9. The dashed blue line indicates the domain decomposition over several processes. To update the particle distributions at the shown lattice point during the propagate step requires input from 9 points, out of which 5 are owned by 3 other processes.

where f_α is the particle distribution in the α-th direction and $f^{(eq)}$ is the equilibrium distribution. For evaluating this expression one usually proceeds in two steps: First, the RHS is computed (this is called the *collision step*), which is followed by the assignment to the LHS (*propagate step*). The last step does not involve any floating-point operations and thus its performance completely depends on the memory performance.

Let use consider the collision step in more detail. There are different formulations of LBM. To apply the methodology introduced in section 2 we use the relatively simple, 2-dimensional D2Q9 model (see Fig. 5 for a graphical representation). To compute the RHS of Eq. (11) at a single location we have to load all f_α and store the same number of intermediate results. For a lattice of size L^2 we thus have an information exchange between processor and external memory of $I_{\text{mem}} = L^2 \cdot 18 \cdot 8\,\text{Byte}$ when this calculation is done in double-precision. The arithmetic computations require almost 80 floating-point operations: $I_{\text{fp}} = L^2 \cdot 78\,\text{Flop}$. Therefore, the Arithmetic Intensity AI$= I_{\text{fp}}/I_{\text{mem}} \simeq 0.6\,\text{Flop/Byte}$. From Fig. 3 we can conclude that also the collision step is memory bound.

For more complex and more frequently used LBM models the AI is much larger. This allows for very efficient implementations of LBM based applications on both processors as well as accelerators. For instance, in case of the D2Q37 model the number of values which need to be loaded and stored only increases to 37 while the number of floating-point operations increases to almost 5,772, leading to an AI of almost 10. The roofline model indicates that for compute devices similar to the ones considered in this lecture, that the application is compute bound. While according to this naive version the roofline model maximum compute efficiency is attainable, this is in practice not achieved. One reason is that we argue purely on basis of nominal performance numbers B_{fp} and B_{mem}. Furthermore, other performance limiting aspects have not been discussed in this lecture. Bortolotti et al.[20] demonstrated that a floating-point performance efficiency of 70% and 62% could be achieved on Intel Xeon processors and NVIDIA K20x GPUs, respectively.

Parallelization of LBM applications is relatively easy. Depending on the size of the problem that is considered, the inherent data-parallelism is large. During the collision and

propagate step each location can be updated independently. The propagate step, however, requires the update of a halo as some of the input data is owned by other processes (see Fig. 5). Assuming a regular d-dimensional lattice (with $d = 2, 3$) of size L^d, parallelization involves a domain decomposition in d' dimensions (with $d' \leq d$). For simplicity let us assume $d' = d$ and a local domain size l^d with $l = L/N^{1/d}$. For the information exchange between node and the network we thus obtain:

$$I_{\text{net}} = b \cdot 2 \cdot d \cdot l^{d-1}, \tag{12}$$

where b is the amount of memory needed to store the distribution functions needed for a single location. In case of the D2Q9 model we find $b = 3 \cdot 8\,\text{Byte}$, as one easily see from Fig. 5. To avoid network communication becoming a bottleneck requires

$$\frac{I_{\text{net}}}{B_{\text{net}}} \ll \max\left(\frac{I_{\text{fp}}}{B_{\text{fp}}}, \frac{I_{\text{mem}}}{B_{\text{mem}}}\right). \tag{13}$$

(Here we again assumed that we can ignore start-up latencies.) Note that for fixed hardware parameters B_{net}, B_{fp} and B_{mem} as well as fixed problem size L^d the LHS and RHS scale differently with the number of nodes, namely $N^{-(d-1)/d}$ and N^{-1}, respectively. This means that by reducing N the network bandwidth requirements can be relaxed.

Because of the typically high AI, the collision step can be efficiently implemented on accelerator devices like GPUs. As the required total amount of memory for this application is relatively small, all data can be kept in the memory of the accelerator device. Therefore the link connecting this device and the processor is not expected to be a bottleneck for this type of applications.

8 Modern Supercomputer Architectures

In the previous sections we discussed several key components of modern HPC architectures. We finish with an overview on two state-of-the-art HPC architectures and specific installations based on these architectures.

The IBM Blue Gene/Q architecture became available in 2012.[21–23] As of November 2014, 4 of the top 10 systems of the Top500[3] List are based on this architecture.

For this architecture a processor is used that had been specifically designed for this architecture. Each processor contains 17 IBM A2 cores, 16 are used for executing user programs while the remaining one runs a slim-lined operating system. The cores are relatively simple, with instructions being executed in order. They support 4-way Simultaneous Multi-Threading (SMT). The processor implements the POWER Instruction Set Architecture (ISA) extended by the Quad Processing eXtension (QPX) ISA. The QPX instructions are executed by a co-processor unit which comprises a 4-way SIMD unit. At each clock cycle this co-processor can execute 4 multiply-add operations. Running at a core clock frequency of 1.6 GHz the processor's peak performance thus is 204.8 GFlop/s.

A key differentiating feature of the Blue Gene/Q processor architecture is the network that has been integrated on the chip. There are 11 links on each processor, out of which 10 are used to connect the processor with its 10 neighbours within a 5-dimensional torus. The 11-th link can be used to connect a compute node to an I/O node, through which communication with the data center network and external storage is realised.

A Blue Gene/Q processor and 16 GByte of memory are integrated on a small node card. 1024 of these node cards can be mounted in a single rack. This very dense design is possible because the racks are liquid cooled, such that there is no need to leave space for airflow.

One of the largest existing Blue Gene/Q installations is hosted by the Jülich Supercomputing Centre. It comprises 28 racks. Selected hardware parameters are provided in Table 3.

The CRAY XC30 architecture differs from the one of Blue Gene/Q in different aspects. The node architecture is based on a commodity processor, namely Intel Xeon, which is based on the x86 ISA. Additionally, accelerator devices can be integrated.

The network is based on the Aries network processor developed by CRAY, which comprises a 48-port router.[24] 8 of its ports are used to attach 4 compute nodes to each router. Another 30 ports are used to connect 96 routers in a 2-dimensional all-to-all structure, while the remaining 10 ports are used for the top-level all-to-all interconnect. This network topology would in principle allow to interconnect 92,544 nodes such that each node could reach any other node in at most 5 network hops.

As of November 2014 the largest XC30 system is installed at the Swiss national supercomputing centre CSCS. Each node of that system comprises an Intel Xeon E5-2670 processor plus an NVIDIA K20x GPU. Selected hardware parameters are again provided in Table 3.

	JUQUEEN	Piz Daint
Number of nodes	28,672	5,272
B_{fp} [MFlop/cycle]	3.7	9.8
B_{fp} [PFlop/s]	5.9	7.8
B_{mem} [PByte/s]	1.2	1.6
C_{mem} [TiByte]	448	196
Power [MWatt]	2.3	2.3

Table 3: Key hardware parameters of the Blue Gene/Q system JUQUEEN installed at Jülich Supercomputing Centre (Germany) and Piz Daint, a CRAY XC30 at CSCS (Switzerland). The listed power consumption was measured during execution of the High-Performance LINPACK benchmark. For typical applications running on these systems power consumption is lower.

9 Conclusions and Outlook

The floating-point capabilities of today's HPC systems result from extreme parallelism. Efficient exploitation of these systems requires concurrent execution of $O(10^6)$ floating-point operations. Around 2017 and after 2020 it is expected that the next generation of parallel computer architectures systems will become available with a peak compute performance of $O(100)$ PFlop/s and $O(1)$ EFlop/s, respectively. This will, however, only be possible by yet another significant increase of parallelism.

Being able to utilize the performance of such systems will require significant efforts from developers of algorithms and applications. While challenging in use, high-end HPC

systems, however, open significant opportunities for scientific research and the creation of new scientific knowledge and results, as shown in the other lectures of this school.

To learn more about computer architectures in general and the use of high-performance computing systems for scientific applications we recommend the textbooks of Hennessy and Patterson[25], Ananth Grama et al.,[26] as well as Hager and Wellein[27] for further reading.

References

1. Martyn Guest et al., "PRACE Scientific Case for HPC in Europe 2012 – 2020", 2012.
2. Steve Ashby et al., "The Opportunities and Challenges of Exascale Computing", 2010.
3. *Top500 List*, http://www.top500.org.
4. Gordon E. Moore, *Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff.*, Solid-State Circuits Society Newsletter, IEEE, **11**, no. 5, 33–35, 2006.
5. J. A. Ang, R. F. Barrett, R. E. Benner, D. Burke, C. Chan, J. Cook, D. Donofrio, S. D. Hammond, K. S. Hemmert, S. M. Kelly, H. Le, V. J. Leung, D. R. Resnick, A. F. Rodrigues, J. Shalf, D. Stark, D. Unat, and N. J. Wright, *Abstract Machine Models and Proxy Architectures for Exascale Computing*, in: Proceedings of the 1st International Workshop on Hardware-Software Co-Design for High Performance Computing, Co-HPC '14, pp. 25–32, IEEE Press, Piscataway, NJ, USA, 2014.
6. Torsten Hoefler, William Gropp, William Kramer, and Marc Snir, *Performance Modeling for Systematic Performance Tuning*, in: State of the Practice Reports, SC '11, pp. 6:1–6:12, ACM, New York, NY, USA, 2011.
7. Gianfranco Bilardi, Andrea Pietracaprina, Geppino Pucci, Fabio Schifano, and Raffaele Tripiccione, "The potential of on-chip multiprocessing for qcd machines", in: High Performance Computing - HiPC 2005, DavidA. Bader, Manish Parashar, Varadarajan Sridhar, and ViktorK. Prasanna, (Eds.), vol. 3769 of *Lecture Notes in Computer Science*, pp. 386–397. Springer Berlin Heidelberg, 2005.
8. Mark Harris, *Mapping Computational Concepts to GPUs*, in: ACM SIGGRAPH 2005 Courses, SIGGRAPH '05, ACM, New York, NY, USA, 2005.
9. Samuel Williams, Andrew Waterman, and David Patterson, *Roofline: An Insightful Visual Performance Model for Multicore Architectures*, Commun. ACM, **52**, no. 4, 65–76, Apr. 2009.
10. R.H. Dennard, F.H. Gaensslen, Hwa-Nien YU, V. Leo Rideovt, Ernest Bassous, and Andre R. Leblanc, *Design of ion-implanted MOSFET's with very small physical dimensions*, Solid-State Circuits Society Newsletter, IEEE, **12**, no. 1, 38–50, Winter 2007.
11. E.J. Nowak, *Maintaining the benefits of CMOS scaling when scaling bogs down*, IBM Journal of Research and Development, **46**, no. 2.3, 169–180, March 2002.
12. L. Chang, D.J. Frank, R.K. Montoye, S.J. Koester, B.L. Ji, P.W. Coteus, R.H. Dennard, and W. Haensch, *Practical Strategies for Power-Efficient Computing Technologies*, Proceedings of the IEEE, **98**, no. 2, 215–236, Feb 2010.
13. B.S. Landman and Roy L. Russo, *On a Pin Versus Block Relationship For Partitions of Logic Graphs*, Computers, IEEE Transactions on, **C-20**, no. 12, 1469–1479, Dec 1971.

14. M.Y. Lanzerotti, G. Fiorenza, and R.A. Rand, *Microminiature packaging and integrated circuitry: The work of E. F. Rent, with an application to on-chip interconnection requirements*, IBM Journal of Research and Development, **49**, no. 4.5, 777–803, July 2005.
15. P. Ruch, T. Brunschwiler, W. Escher, S. Paredes, and B. Michel, *Toward five-dimensional scaling: How density improves efficiency in future computers*, IBM Journal of Research and Development, **55**, no. 5, 15:1–15:13, Sept 2011.
16. Gyozo I. Egri, Zoltán Fodor, Christian Hoelbling, Sándor D. Katz, Dániel Nógrádi, and Kálmán K. Szabó, *Lattice QCD as a video game*, Computer Physics Communications, **177**, no. 8, 631–639, 2007.
17. C.E. Leiserson, *Fat-trees: Universal networks for hardware-efficient supercomputing*, Computers, IEEE Transactions on, **C-34**, no. 10, 892–901, Oct 1985.
18. J. Kim, W.J. Dally, S. Scott, and D. Abts, *Technology-Driven, Highly-Scalable Dragonfly Topology*, in: Computer Architecture, 2008. ISCA '08. 35th International Symposium on, pp. 77–88, June 2008.
19. Christian Feichtinger, Jan Götz, Stefan Donath, Klaus Iglberger, and Ulrich Rüde, "WaLBerla: Exploiting massively parallel systems for Lattice Boltzmann simulations", in: Parallel Computing, Roman Trobec, Marian Vajtersic, and Peter Zinterhof, (Eds.), pp. 241–260. Springer London, 2009.
20. G. Bortolotti, M. Caberletti, G. Crimi, A. Ferraro, F. Giacomini, M. Manzali, G. Maron, M. Pivanti, D. Salomoni, S. F. Schifano, R. Tripiccione, and M. Zanella, *Computing on Knights and Kepler Architectures*, Journal of Physics: Conference Series, **513**, no. 5, 052032, 2014.
21. K. Sugavanam, C.-Y. Cher, J. A. Gunnels, R. A. Haring, P. Heidelberger, H. M. Jacobson, M. K. McManus, D. P. Paulsen, D. L. Satterfield, Y. Sugawara, and R. Walkup, *Design for low power and power management in IBM Blue Gene/Q*, IBM Journal of Research and Development, **57**, no. 1/2, 3:1–3:11, Jan 2013.
22. IBM Blue Gene team, *Design of the IBM Blue Gene/Q Compute chip*, IBM Journal of Research and Development, **57**, no. 1/2, 1:1–1:13, Jan 2013.
23. P. W. Coteus, S. A. Hall, T. Takken, R. A. Rand, S. Tian, G. V. Kopcsay, R. Bickford, F. P. Giordano, C. M. Marroquin, and M. J. Jeanson, *Packaging the IBM Blue Gene/Q supercomputer*, IBM Journal of Research and Development, **57**, no. 1/2, 2:1–2:13, Jan 2013.
24. G. Faanes, A. Bataineh, D. Roweth, T. Court, E. Froese, B. Alverson, T. Johnson, J. Kopnick, M. Higgins, and J. Reinhard, *Cray Cascade: A scalable HPC system based on a Dragonfly network*, in: High Performance Computing, Networking, Storage and Analysis (SC), 2012 International Conference for, pp. 1–9, Nov 2012.
25. John L. Hennessy and David A. Patterson, *Computer Architecture, Fifth Edition: A Quantitative Approach*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 5th edition, 2011.
26. Ananth Grama, Anshul Gupta, George Karypis, and Vipin Kumar, *Introduction to Parallel Computing*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 2002.
27. Georg Hager and Gerhard Wellein, *Introduction to High Performance Computing for Scientists and Engineers*, CRC Press, Inc., Boca Raton, FL, USA, 1st edition, 2010.

Schriften des Forschungszentrums Jülich
IAS Series

Band / Volume 16
Design and Applications of an Interoperability Reference Model for Production e-Science Infrastructures
by M. Riedel (2013), x, 270 pages
ISBN: 978-3-89336-861-7
URN: urn:nbn:de:0001-2013031903

Band / Volume 17
Route Choice Modelling and Runtime Optimisation for Simulation of Building Evacuation
by A. U. Kemloh Wagoum (2013), xviii, 122 pages
ISBN: 978-3-89336-865-5
URN: urn:nbn:de:0001-2013032608

Band / Volume 18
Dynamik von Personenströmen in Sportstadien
by S. Burghardt (2013), xi, 115 pages
ISBN: 978-3-89336-879-2
URN: urn:nbn:de:0001-2013060504

Band / Volume 19
Multiscale Modelling Methods for Applications in Materials Science
by I. Kondov, G. Sutmann (2013), 326 pages
ISBN: 978-3-89336-899-0
URN: urn:nbn:de:0001-2013090204

Band / Volume 20
High-resolution Simulations of Strongly Coupled Coulomb Systems with a Parallel Tree Code
by M. Winkel (2013), xvii, 196 pages
ISBN: 978-3-89336-901-0
URN: urn:nbn:de:0001-2013091802

Band / Volume 21
UNICORE Summit 2013
Proceedings, 18th June 2013 | Leipzig, Germany
edited by V. Huber, R. Müller-Pfefferkorn, M. Romberg (2013), iii, 94 pages
ISBN: 978-3-89336-910-2
URN: urn:nbn:de:0001-2013102109

Band / Volume 22
Three-dimensional Solute Transport Modeling in Coupled Soil and Plant Root Systems
by N. Schröder (2013), xii, 126 pages
ISBN: 978-3-89336-923-2
URN: urn:nbn:de:0001-2013112209

Schriften des Forschungszentrums Jülich
IAS Series

Band / Volume 23
Characterizing Load and Communication Imbalance in Parallel Applications
by D. Böhme (2014), xv, 111 pages
ISBN: 978-3-89336-940-9
URN: urn:nbn:de:0001-2014012708

Band / Volume 24
Automated Optimization Methods for Scientific Workflows in e-Science Infrastructures
by S. Holl (2014), xvi, 182 pages
ISBN: 978-3-89336-949-2
URN: urn:nbn:de:0001-2014022000

Band / Volume 25
Numerical simulation of gas-induced orbital decay of binary systems in young clusters
by A. C. Korntreff (2014), 98 pages
ISBN: 978-3-89336-979-9
URN: urn:nbn:de:0001-2014072202

Band / Volume 26
UNICORE Summit 2014
Proceedings, 24th June 2014 | Leipzig, Germany
edited by V. Huber, R. Müller-Pfefferkorn, M. Romberg (2014), iii, 60 pages
ISBN: 978-3-95806-004-3
URN: urn:nbn:de:0001-2014111408

Band / Volume 27
Automatische Erfassung präziser Trajektorien in Personenströmen hoher Dichte
by M. Boltes (2015), xii, 308 pages
ISBN: 978-3-95806-025-8
URN: urn:nbn:de:0001-2015011609

Band / Volume 28
Computational Trends in Solvation and Transport in Liquids
edited by Godehard Sutmann, Johannes Grotendorst, Gerhard Gompper, Dominik Marx (2015), ?? pages
ISBN: 978-3-95806-030-2
URN: urn:nbn:de:0001-2015020300

Weitere ***Schriften des Verlags im Forschungszentrum Jülich*** unter
http://wwwzb1.fz-juelich.de/verlagextern1/index.asp

The majority of chemical reactions including many important industrial processes and virtually all biological activities take place within a liquid environment. Solvents, of which water is certainly the most important, are able to “solvate” molecules, thereby transferring these as “solutes” into the liquid state. Transport processes and solute-solute interactions in the solvent are then supporting structure formation, self-organization or chemical reactions. Solvents are not only able to provide a liquid phase for simple chemical reagents and the much more complex proteins; they have the additional ability to wet extended surfaces such as lipid membranes or metal electrodes, thereby creating interfaces. An in-depth understanding of solvation at a fundamental level of chemistry, physics and engineering is essential to enable major advances in key technologies for environmentally friendly technologies, e.g. to reduce pollution, to increase energy efficiency or to prevent corrosion to name but a few challenges to our modern day society. In biophysics and life sciences, water is the most important and dominant solvent, providing the basic environment for the complexity of life. Therefore, an understanding of solvation is crucial to unravel biological function in a comprehensive way.

The Lecture Notes contain the current state-of-the-art methods to treat solvation and transport on different levels of resolution. Topics include ab initio methods, atomistic and mesoscale methods for modeling accurately the solute-solvent interaction and an efficient treatment of the solvent on a mesoscopic level. Recent advances in mathematical techniques are introduced, which are fundamental for efficient treatment of solute-solvent systems. Recent trends and future directions in computational science are addressed to provide a perspective for software development and computer architectures.

This publication was edited at the Jülich Supercomputer Centre (JSC) which is an integral part of the Institute for Advanced Simulation (IAS). The IAS combines the Jülich Simulation sciences and the supercomputing facility in on organizational unit. It includes those parts of the scientific institutes at Forschungszentrum Jülich which use simulation on supercomputers as their main research methodology.

IAS Series
Volume 28
ISBN 978-3-95806-030-2

JÜLICH
FORSCHUNGSZENTRUM